



VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA

EKONOMICKÁ FAKULTA

KATEDRA SYSTÉMOVÉHO INŽENÝRSTVÍ

Návrh a implementace řešení reportingu pro management oddělení IT společnosti

Design and Implementation of a Reporting Solution for the Management  
of an IT Company Department

Student: Bc. Kateřina Martynková

Vedoucí diplomové práce: Ing. Radek Němec, Ph.D.

Ostrava 2016

## Zadání diplomové práce

Student: **Bc. Kateřina Martynková**

Studijní program: N6209 Systémové inženýrství a informatika

Studijní obor: 6209T025 Systémové inženýrství a informatika

Téma: Návrh a implementace řešení reportingu pro management oddělení IT společnosti  
Design and Implementation of a Reporting Solution for the Management of an IT Company Department

Jazyk vypracování: čeština

Zásady pro vypracování:

1. Úvod
2. Teoretická východiska související s životním cyklem vývoje řešení reportingu
3. Analýza současného stavu zpracování a poskytování reportů v oddělení IT společnosti
4. Návrh řešení a implementace
5. Závěr

Seznam použité literatury

Seznam zkratk

Prohlášení o využití výsledků diplomové práce

Seznam příloh

Přílohy

Seznam doporučené odborné literatury:

BALLARD, Chuck a kol. *Dimensional Modeling: In a Business Intelligence Environment*. San Jose: IBM Redbooks, 2006. ISBN 0738496448.

POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. *Business intelligence v podnikové praxi*. Praha: Professional Publishing, 2012. ISBN 978-80-7431-065-2.

NĚMEC, Radek. *Principy projektování a implementace systémů Business Intelligence*. Ostrava: VŠB-TU Ostrava, 2014. ISBN 978-80-248-3452-8.

Formální náležitosti a rozsah diplomové práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí diplomové práce: **Ing. Radek Němec, Ph.D.**

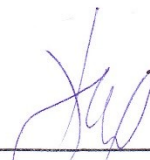
Datum zadání: 20.11.2015

Datum odevzdání: 22.04.2016



---

doc. Ing. Jana Hančlová, CSc.  
vedoucí katedry



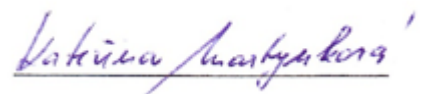
---

prof. Dr. Ing. Dana Dluhošová  
děkanka fakulty

**Místopřísežné prohlášení o samostatném vypracování diplomové práce**

„Prohlašuji, že jsem celou práci, včetně všech příloh, vypracovala samostatně“.

Zároveň bych na tomto místě chtěla poděkovat svému vedoucímu práce Ing. Radkovi Němcovi, Ph.D. za vedení diplomové práce, připomínky a rady. Dále také celé své rodině za možnost studovat a příteli za velkou podporu.



Bc. Kateřina Martynková

Datum a odevzdání diplomové práce: 22. dubna 2016

# Obsah

1.	Úvod .....	5
2.	Teoretická východiska spojená s životním cyklem vývoje řešením reportingu.....	7
2.1.	Business Intelligence .....	7
2.1.1.	Vymezení pojmu Business Intelligence .....	7
2.1.2.	Životní cyklus vývoje systému BI.....	8
2.2.	Datové modelování .....	9
2.2.1.	Sémantický datový model .....	10
2.2.2.	Konceptuální datový model.....	11
2.2.3.	Logický datový model.....	12
2.2.4.	Shrnutí .....	12
2.3.	Dimenzionální modelování .....	13
2.3.1.	Princip dimenzionálního modelování.....	13
2.3.2.	Životní cyklus dimenzionálního modelování .....	14
2.3.3.	Specifikace podnikového procesu a požadavků .....	16
2.3.4.	Vymezení granularity .....	18
2.3.5.	Identifikace dimenzí .....	20
2.3.6.	Identifikace faktů.....	23
2.3.7.	Fyzický návrh dimenzionálního modelu .....	24
2.3.8.	Shrnutí .....	26
2.4.	Reprezentace a uložení dat multidimenzionálního modelu .....	27
2.4.1.	Relační databáze .....	27
2.4.2.	Multidimenzionální databáze .....	28
2.4.3.	Shrnutí .....	30
2.5.	Komponenty pro analýzu a reporting.....	30
2.5.1.	Multidimenzionální analýzy .....	30
2.5.2.	Analytické a reportingové nástroje.....	31
2.5.3.	Dashboardy na principu User Experience .....	32
2.5.4.	Shrnutí .....	34
3.	Analýza současného stavu zpracování a poskytování reportů v oddělení IT společnosti.....	35
3.1.	Popis firmy a firemního prostředí .....	35
3.1.1.	Základní charakteristika firmy .....	35
3.1.2.	Organizační struktura firmy .....	36
3.1.3.	Procesy ve firmě.....	37

3.2.	Popis oddělení <i>Customer Service Centre</i> .....	38
3.2.1.	Procesy v oddělení <i>Customer Service Centre</i> .....	39
3.3.	Proces Incident Management .....	39
3.3.1.	Životní cyklus řešení incidentu .....	41
3.4.	Současný stav zpracování a poskytování reportů .....	42
3.4.1.	Role v reportingu SD .....	43
3.4.2.	Poskytované metriky a reporty .....	44
3.4.3.	Nástroje pro zpracování reportů .....	49
3.4.4.	Procesy vytváření metrik a lidské zdroje .....	51
3.4.5.	Zdroje pro reporty .....	51
3.4.6.	Možnosti dodávání reportů / metrik .....	52
3.4.7.	Shrnutí .....	53
4.	Návrh řešení a implementace .....	54
4.1.	Dimenzionální modelování .....	54
4.1.1.	Zmapování podnikových procesů a požadavků .....	54
4.1.2.	Stanovení granularity .....	58
4.1.3.	Identifikace dimenzí .....	59
4.1.4.	Identifikace faktů .....	66
4.1.5.	Fyzický návrh .....	68
4.2.	Implementace datových pump .....	69
4.2.1.	Určení zdrojových a cílových datových struktur .....	70
4.2.2.	Kontrola dat .....	71
4.2.3.	Návrh rutin .....	71
4.2.4.	Základní transformace .....	75
4.3.	Návrh výstupů a jejich prezentace .....	75
4.3.1.	Sběr požadavků výstupů .....	75
4.3.2.	Návrh výstupů .....	77
4.3.3.	Zhodnocení výstupů .....	79
4.3.4.	Zveřejnění výstupů .....	83
5.	Závěr .....	84
	Seznam použité literatury .....	86
	Seznam zkratk .....	88
	Prohlášení o využití výsledků diplomové práce .....	
	Seznam příloh .....	

# 1. Úvod

Firmy svá rozhodnutí o budoucím vývoji podniku mohou provádět na základě výsledků firemní aktivity. Nejsou-li výsledky relevantní, správné, či včas a strategicky přístupné, rozhodnutí může být aplikované nepřesně nebo pozdě, což může mít za následek negativní dopad na podnikání. V dnešním světě mnoha technologií, nástrojů a chytrých zařízení je možné stav firmy sledovat v reálném čase a manažeři tak mohou rychle jednat. Důležité také ale je, aby firmy měly vhodně vybrané a nastavené metriky i klíčové indikátory výkonnosti, mít je strategicky dostupné, mít dobře nastavené procesy zpracování poskytovaných výsledků a mít zajištěné jejich řízení a kontrolování.

Cílem diplomové práce je v oddělení nabízející IT služby navrhnout reportingovou platformu pro zpracovávání reportů a dashboardů, která bude poskytovat přehlednější výsledky oddělení. Účelem platformy je, pomocí standardizace KPI, metrik, reportů, procesu jejich vytváření, multidimenzionálního pohledu na data a interaktivního reportingu, zefektivnit procesy v daném oddělení a zajistit tak lepší kontrolu nad kvalitou poskytované služby.

Diplomová práce je strukturována do pěti hlavních kapitol – úvod, teoretická východiska, analýza současného stavu, návrh řešení a implementace a závěr.

V první kapitole *Úvod* je vytyčen cíl diplomové práce a je stručně popsán obsah jednotlivých kapitol.

Druhá kapitola *Teoretická východiska spojená s životním cyklem vývoje řešením reportingu* je zaměřena na vyličení důležitých teoretických poznatků potřebných k vypracování praktické části práce. Nejdřív je věnována pozornost pojmu Business Intelligence a jeho životnímu cyklus. Dále je popsáno datové modelování jako nezbytná součást při navrhování všech systémů. Největší pozornost je zaměřena na dimenzionální modelování, jakožto hlavní pilíř diplomové práce. V této kapitole jsou čtenáři seznámeni s životním cyklem R. Kimballa – principy pro specifikaci procesů a požadavků, vymezení granularity, identifikaci dimenzí a faktů. Popsány jsou možnosti fyzického návrhu u dimenzionálního modelování. Jsou demonstrovány příležitosti uložení dimenzionálně uspořádaných dat. V poslední části jsou představeny komponenty pro analýzu a reporting jako multidimenzionální techniky, nebo návrhy dashboardů na principech User Experience.



Ve třetí kapitole *Analýza současného stavu zpracování a poskytování reportů v oddělení IT společnosti* je popsána vybraná firma včetně organizační struktury a procesů. Podrobnější popis je zaměřen na samotné IT oddělení společně s procesy, pro které je navrhován dimenzionální model. Jsou vypsány doposud poskytované reporty a metriky vzhledem k jejich principu a omezení. Jsou vytýčeny role v reportingu, používané nástroje, procesy zpracování a lidské zdroje, zdroje pro reporty a možnosti ukládání a sdílení.

Ve čtvrté kapitole *Návrh řešení a implementace* jde o praktickou realizaci diplomové práce, která je rozdělena do tří částí – dimenzionální modelování, implementace datových pump a návrh výstupních dashboardů. V dimenzionálním modelování je nejdříve vyspecifikován zkoumaný proces, požadavky uživatelů, zdrojové a cílové struktury. Dále je stanovena granularita v tabulkách faktů, jsou identifikovány tabulky dimenzí a faktů a je proveden fyzický návrh. Následně jsou navrhnuty datové pumpy a implementovány datové rutiny pro nahrání dat do dimenzionálního modelu. Nakonec jsou navrženy výsledné dashboardy, jsou zhodnoceny stanovené požadavky a je otestována funkčnost dashboardů.

V páté kapitole *Závěr* je provedeno celkové zhodnocení práce, od analýzy společnosti až po samotný návrh řešení a implementaci. Jsou vytýčeny přínosy z navrhnutého řešení a jsou zmíněna očekávání a kroky do budoucna.

## **2. Teoretická východiska spojená s životním cyklem vývoje řešením reportingu**

V této kapitole jsou popsány hlavní části spojené s vývojem řešení reportingu. Ve zkratce je vymezena souvislost s Business Intelligence. Podrobněji je zaměřeno na datové modelování, dimenzionální modelování, reprezentaci a uložení dat v multidimenzionálním modelu, analýzu a reporting.

### **2.1. Business Intelligence**

Hlavním východiskem této podkapitoly je vymezit co je to Business Intelligence, její architektura a životní cyklus vývoje systému Business Intelligence.

#### **2.1.1. Vymezení pojmu Business Intelligence**

Již od vzniku prvních firem bylo velmi důležité, aby tyto firmy efektivně podporovaly svou činnost. Firmy musely zaznamenávat veškerou aktivitu a výstupy. Tato data musela být analyzována a na základě analýz se mohlo provádět plánování a rozhodování o budoucích aktivitách podniku. Neexistovaly žádné stroje, vše se muselo provádět jen pomocí tužky a papíru. Postupem času a vývojem prvních počítačů mohly firmy využít tuto technologii pro ukládání dat či informací a také tyto stroje mohly využít k výše zmíněným krokům.

V současném rychle se rozvíjejícím měnícím se prostředí a konkurenčním světě podnikání musí firmy využít jiných technologií, nástrojů, metod, aplikací a přístupů, aby byly schopné co nejrychleji a optimálně reagovat na situace v jejich podniku a na trhu a staly se odolné vůči konkurenci. Pro zpracování dat, jejich prezentaci a následné zefektivnění procesů podniku mají dnes firmy možnost výběru z kvalitních informačních technologií a návrhu přesně odpovídajícímu informačnímu systému.

Systém s takovými možnostmi byl poprvé v roce 1989 označen jako *Business Intelligence* (dále BI), a to analytikem Howardem. J. Dresnerem ze společnosti Gartner. Pomocí systému BI je možné v podniku nalézt oblasti, které přispívají k růstu organizace nebo, naopak oblasti, jež podniku škodí. Pro pojem BI existuje mnoho definic. Z řad odborníků se jedná o tyto:

Pour, Maryška a Novotný (2012) definují BI jako sadu procesů, know-how, aplikací a technologií, jejichž cílem je účinně a účelně podporovat řídicí aktivity ve firmě. Podporují

analytické, plánovací a rozhodovací činnosti organizací na všech úrovních a ve všech oblastech podnikového řízení, tj. prodeje, nákupu, marketingu, finančního řízení controllingu, majetku, řízení lidských zdrojů, výroby a dalších.

Gartner (2015) popisuje BI jako zastřešující termín, který zahrnuje aplikace, infrastrukturu, nástroje a osvědčené postupy, jež umožňují přístup a analýzu informací, které mají zvýšit a optimalizovat rozhodnutí a výkonnost.

Architekturu BI je možné rozdělit do tří částí, a to produkční databáze, samotné komponenty BI a koncoví uživatelé používající řešení BI.

**Architektura BI obsahuje:**

- produkční databáze,
- komponenty datové transformace - ETL (Extract, Transform, Load), ELT (Extract, Load, Transform), EAI (Enterprise Application Integration),
- databázové komponenty - dočasná úložiště (Data Staging Area – DSA), datové sklady (Data Warehouse – DW), datová tržiště (Data Marts – DM), operativní datová úložiště (Operation Data Store – ODS),
- analytické komponenty – OLAP, reporting,
- dolování data,
- datová kvalita a metadata.

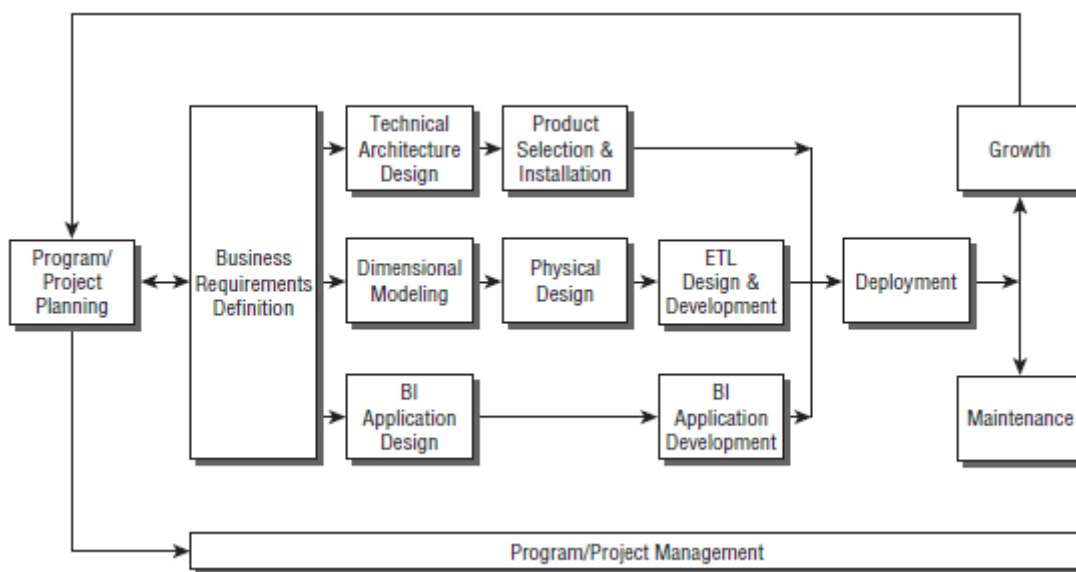
### **2.1.2. Životní cyklus vývoje systému BI**

U budování informačních systémů je nutné, aby byl vývoj veden jako projekt, neboť je systém složen z mnoha činností, komponent a vazeb. Projekt musí z obecného hlediska projít následujícími fázemi:

- plánováním a přípravou,
- analýzou a návrhem,
- implementací,
- zavedením do provozu, migrací a distribucí,
- provozem,
- rozvojem a optimalizací.

Výše zmíněné fáze jsou dále modifikovány pro systémy BI. Rozdílů jsou rovněž ve firmou zvolené metodice<sup>1</sup>, kterých je celá řada.

V další části textu se ve velké míře bude vycházet z metodiky R. Kimballa. Jedná se o životní cyklus vývoje systému BI dle Kimballa a Rosse (Kimball a Ross, 2013), který je vidět na obrázku Obr 2.1.



Obr. 2.1 Diagram životního cyklu vývoje systému BI (Kimball a Ross, 2013)

Jelikož je dimenzionální modelování, jak i z Obr. 2.1 vyplývá, součástí BI, byl zde tento pojem charakterizován pro jasnější vytýčení směru práce vzhledem ke všem oblastem informačních technologií. Další části práce jsou zaměřeny na dimenzionální modelování, kde budou popsány jednotlivé fáze cyklu, jejich propojení a návaznost. Pozornost bude soustředěna také na metody<sup>2</sup>, jejich techniky a prostředky, které jsou pro danou fázi typické.

## 2.2. Datové modelování

Při návrhu informačního systému, jak už pro BI, tak i pro jakékoli další systémy, by analytici měli abstrahovat firemní procesy od reality. Měli by realitu vystihnout pomocí modelů a graficky zachytit požadovaný výstup. Většině lidí se může tento krok jevit jako nevýznamný a zbytečný, ovšem opak je pravdou. Výstupem datového modelování jsou analytiky zpracovány

<sup>1</sup> Metodika - je souhrn doporučených praktik a postupů, které by měla zahrnovat motivy, poskytovatele i strategický náhled (Kaluža, 2010).

<sup>2</sup> Metoda – určuje, co je třeba udělat v určité etapě vývoje životního cyklu (Němec, 2014).

modely na třech úrovních koncepce, které slouží jako vstupní materiály pro práci programátorů. Funkcí programátorů v rámci projektu není zjišťování entit, jejich atributů a vazeb, ale pouze kódování. Tudíž kdyby došlo k vypuštění tohoto kroku, neměli by programátoři z čeho čerpat.

Dle Kaluži (2010) je nutné rozdělit modelování do tří úrovní, tzv. *tříúrovňová koncepce* datového modelování, která je složena ze *sémantické, konceptuální a logické úrovně*. První úroveň je sémantická úroveň, jejímž cílem je vymezení výchozí struktury modelu podle identifikovaných požadavků (neboli typů objektů). Na základě sémantické úrovně je konstruována úroveň konceptuální. V konceptuální úrovni jsou pomocí E-R diagramů nebo diagramů tříd graficky znázorněny jednotlivé třídy (entity) a vztahy mezi nimi. Poslední úroveň je zaměřena na logické modelování, které je spjata s již konkrétní databázovou komponentou.

### 2.2.1.Sémantický datový model

Pro zpracování sémantického modelu jsou důležité požadavky uživatelů – přesně vytýčený cíl a rozsah. Z požadavků jsou specifikovány předběžné typy objektů, které jsou pro daný systém charakteristické.

V rámci systému BI se jedná o nalezení podnikových procesů, pro které bude navržen dimenzionální model. V těchto procesech bude stanoven stupeň granularity. Pro něj jsou následně vyhledávány konkrétní dimenze, podle nichž je možné měnit pohled na data. Rovněž jsou pro danou úroveň definována fakta jako zkoumané metriky ve vybraném business procesu. (Ballard, 2006) V tabulce Tab. 2.1 je uveden příklad vymezení zkoumaného procesu, stupeň granularity, návrh dimenzí a metrik.

<b>Zkoumaný proces:</b>	obsluha zákazníka v IT firmě (práce s tikety)
<b>Úroveň podrobnosti:</b>	transakce tiketů, přehledy
<b>Návrh dimenzí:</b>	čas
	incident
	zákazník
	uživatel
	skupina, atd.
<b>Návrh metrik:</b>	počet otevřených / verifikovaných tiketů
	průměrný aktivní čas strávený na tiketu
	průměr hodnocení koncových uživatelů, atd.

Tab. 2.1 Ukázka vybraných dimenzí a metrik (vlastní zpracování)

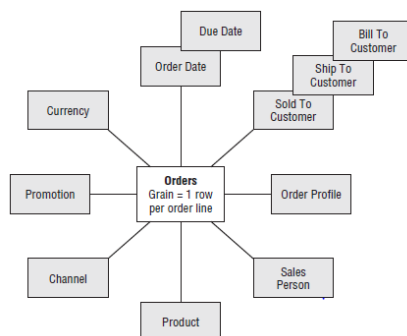
Na hrubě stanoveném modelu bude následně provedena revize dimenzí a metrik. Důvodem mohou být různé názvy dimenzí pro stejný objekt, případně redundance atributů dimenzí, atd. U faktů, jejímž cílem je poskytovat výsledky za stanovenou agregaci, musí být prozkoumáno, zda metriky v agregaci dávají smysl. Ne každá číselná hodnota musí být metrikou. Touto revizí může dojít k odstranění nevyhovujících dimenzí či přidání nových. Stejný postup je uplatněn pro metriky a atributy dimenzí.

## 2.2.2. Konceptuální datový model

Již výše bylo uvedeno, že konceptuální datový model je grafickým zobrazením zjištěných tříd ze sémantického modelu a vazeb mezi nimi. Dle Kaluži (2010) jsou nejužívanějšími metodami vizualizace konceptuálního modelu E-R diagram a diagram tříd (součást metodiky Unified Modeling Language - UML).

Ve vývoji systému BI se jedná o návrh struktury multidimenzionálního modelu. Jde o konkrétní popis dimenzí – jaké jsou jejich atributy, z jakého zdrojového objektu (množiny objektů) je atribut tvořen, jak bude transformován, jakého je dimenze typu, je určen primární klíč dimenze (příp. skupina primárních klíčů), atd. Popsána je rovněž množina faktů – jakého jsou typu, z jakého zdrojového objektu budou metriky vypočteny, jaká je jejich granularita.

Na obrázku Obr. 2.2 je možné vidět konceptuální návrh dimenzionálního modelu. V tomto nákresu lze vidět dimenze business procesu a tabulku faktů. V tabulce faktů je identifikována úroveň podrobnosti. Ke každé dimenzi a množině faktů, jak bylo zmíněno v předchozím odstavci, je např. v tabulkovém procesoru zpracována podrobná tabulka.



Obr. 2.2 Ukázka konceptuálního modelu dimenzí a tabulky faktů (Kimball a Ross, 2013)

Při modelování může nastat situace, kdy jedna dimenze může být součástí více procesů, tudíž více multidimenzionálních modelů. Jde tzv. o *sdílenou dimenzi*, o které Kimball (2002) tvrdí,

že dimenze je sdílená, jestliže dvě dimenze obsahují jednu nebo více atributů se stejným jménem. Pro vizualizaci procesů a jejich odpovídajících dimenzí je proto použit *Kimball Bus Matrix*. Typickou dimenzí, která je často využita ve všech procesech, je dimenze času.

Na obrázku Obr. 2.3 jsou zachyceny procesy firmy a k nim odpovídající dimenze. Např. dimenze *Product* je zastoupena ve všech vytýčených procesech.

BUSINESS PROCESSES	COMMON DIMENSIONS						
	Date	Product	Warehouse	Store	Promotion	Customer	Employee
Issue Purchase Orders	X	X	X				
Receive Warehouse Deliveries	X	X	X				X
Warehouse Inventory	X	X	X				
Receive Store Deliveries	X	X	X	X			X
Store Inventory	X	X		X			
Retail Sales	X	X		X	X	X	X

Obr. 2.3 Ukázka zachycení procesů a jejich dimenzí pomocí Kimball Bus Matrix (Kimball a Ross, 2013)

### 2.2.3. Logický datový model

Na úrovni logického modelování je zajištěna transformace entit do logických struktur. Před samotnou fyzickou realizací dimenzionálního modelu je nutné doladit detaily daného modelu. K modelování je využito softwarového nástroje, v němž je jednotlivým entitám přiřazen datový typ, maximální velikost, integritní omezení – referenční (cizí klíč) a entitní (primární klíč). Tento krok datového modelování je důležitý pro následnou tvorbu rutin datových pump.

Logický datový model může pomoci k ověření funkčnosti modelu vzhledem k jeho požadavkům. Implementace je nezávislá na požadavky fyzického návrhu v databázi. S využitím logického modelu je snazší a jednoznačné nalezení všech podnikatelských entit v systému včetně vazeb. (Ballard, 2006)

### 2.2.4. Shrnutí

Kapitola *Datové modelování* byla do teoretické části práce zahrnuta kvůli její důležitosti, jak bylo i na začátku této kapitoly řečeno. Jeho cílem je stanovit, čeho má být dosaženo, co je

k tomu zapotřebí (např. jaké dimenze, fakta), v jakém jsou entity vztahu, co je důležité nastavit (např. integritní omezení, indexy), atd. Jde o tzv. *resumé* – sjednocení a ukotvení myšlenek celého konceptu projektu a hlavně vstupních materiálů pro samotnou fyzickou realizaci.

Poznatky z této kapitoly budou v praktické části diplomové práce uplatněny v podobě jednotlivých modelů - sémantický, konceptuální a logický model.

## 2.3. Dimenzionální modelování

Smyslem podkapitoly *Dimenzionální modelování* je seznámit čtenáře s principem dimenzionálního modelování, s životním cyklem dimenzionálního modelování a s jeho jednotlivými fázemi - specifikace podnikového procesu a požadavků, vymezení granularity, identifikace dimenzí a faktů a fyzický návrh dimenzionálního modelu.

### 2.3.1. Princip dimenzionálního modelování

Postupné zvyšování objemu dat, které jsou firemními systémy zpracovávány a ukládány, zvyšování nároků manažerů na prezentované informace a zapojení ne-manažerských pozic do procesu rozhodování, bylo důvodem k jinému perspektivnějšímu ukládání a pohlížení na data.

Jde o multidimenzionální pohled na data, kde autorem dimenzionálního modelování je Ralph Kimball. Kimball (2013) o dimenzionálním modelování tvrdí, že *jde o dlouhodobou techniku, jak docílit jednoduchých databází. V dimenzionálním modelování je zahrnuta analýza, návrh struktury a obsahu informačního modelu, který slouží k rozhodovacím potřebám a k efektivnímu řízení.*

Úkolem dimenzionálního modelování je vytvořit takovou logiku uložení a uspořádání dat, kterou bude zajištěna účinná podpora manažerského rozhodování a řízení procesů organizace, a to tak aby rozhodnutí na základě výstupu dimenzionálního modelu bylo smysluplné, intuitivní a rychlé. Výstup modelu by měl být prezentován tou nejtriviálnější formou s co nejkratší dobou odezvy, ale na druhou stranu informace v něm by měly být relevantní a měly by odpovídat realitě podnikových procesů.

Typické pro dimenzionální modelování je, že na údaje, které chce organizace sledovat, je možné pohlížet z různých úhlů pohledů a na různé úrovni detailu. Jinak na údaje budou pohlížet vrcholoví manažeři a jinak manažeři jednotlivých oddělení. Např. vrcholoví manažeři chtějí vidět



výkonnostní výsledky celé organizace, naproti tomu operativní manažeři sledují výsledky svého oddělení, svých podřízených, příp. týmů, atd. Vrcholoví manažeři se v podniku víc zajímají o finanční situaci, nižší management spíš o provozní část.

K dimenzionálnímu modelu jsou využívány tzv. tabulky dimenzí, faktů a jejich vzájemné vztahy. Fakty jsou myšleny ukazatele či metriky organizace, jenž mají být monitorovány. Oblasti, za které mají být ukazatele sledovány, jsou nazývány dimenzemi. Podrobněji budou fakta a dimenze popsány v následujících kapitolách.

Hlavní rozdíl mezi databázemi dimenzionálního modelu a relačními databázemi spočívá v normalizaci dat. V relační databázi jsou data uložena minimálně ve 3. normální formě, kdežto v multidimenzionálním modelu mohou být data nenormalizovaná (nultá normální forma), proto by zpracování velkých objemů dat (zejména při agregacích) nemělo systémy tolik zatěžovat. Rovněž je zajištěno rychlejší nahrávání dat a rychlejší analytické operace nad daty.

### **2.3.2. Životní cyklus dimenzionálního modelování**

Získávání a správné pochopení uživatelských požadavků je velmi obtížný a zdlouhavý proces. Snahou výstupu dimenzionálního modelu je uživateli zajistit informovanost a spokojenost. Existuje proto několik metodologií, jakými kroky toho docílit. V rámci práce je v úvahu brána metodologie Kimballa a Rosse.

Kimball a Ross (2013) uvádí čtyři aktivity životního cyklu dimenzionálního modelování (*identifikace podnikového procesu, deklarace granularity, identifikace dimenzí a identifikace faktů*). Cílem je identifikovat ty fáze a činnosti, na které je potřebné brát ohled při návrhu dimenzionálního modelu. V kombinaci s agilním životním cyklem může jít o velmi efektivní a „příjemný“ návrh modelu, kde spokojenost bude, jak na straně zadavatele, tak na straně zpracovatele. Ve zkratce agilní cyklus spočívá v opakování nevelkých úkolů v krátkých časových úsecích, kde je do samotného procesu aktivně zapojen zákazník, který kontroluje každou dokončenou iteraci a podává připomínky. Jeho náměty mohou být případně v následující iteracích ihned upraveny. Rovněž mohou být opraveny nalezené chyby, což je efektivnější při návrhu modelu jako celku.

#### Čtyři výše zmíněné aktivity:

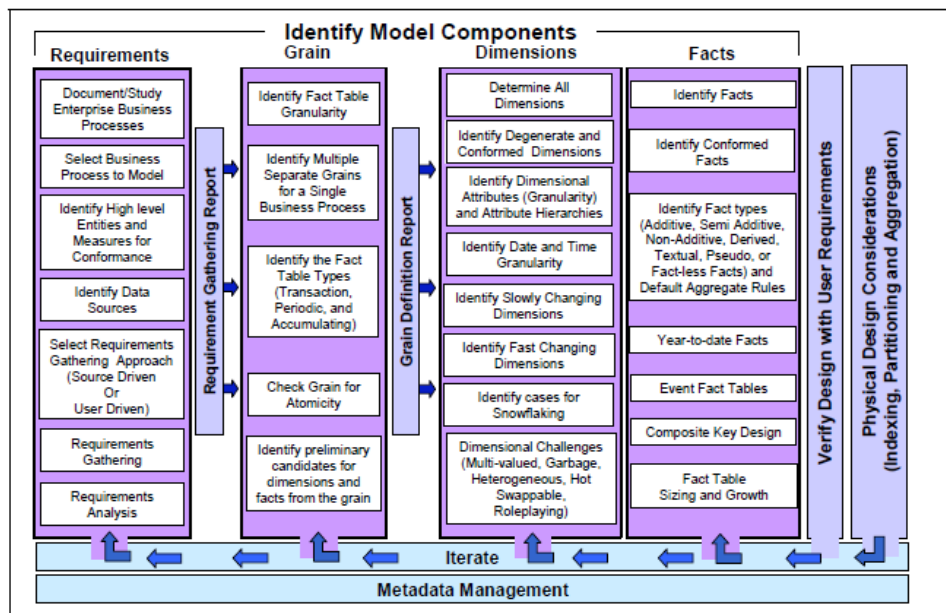
- **Identifikace podnikových procesních požadavků.** Cílem tohoto kroku je výběr podnikového procesu, jenž má být zkoumán. Pro vytýčený proces je navržen samotný dimenzionální model. Na základě provedeného výběru jsou sesbírány konkrétní požadavky od uživatelů (co chtějí v daném procesu zkoumat a za jaké oblasti). Volba daného procesu může být náročná z hlediska množství procesů v podniku. Je vhodné jednotlivým procesům přiřadit priority. Kritéria priorit se mohou podnik od podniku lišit (např. důležitost a význam procesu pro podnik, kvalita dat ve zdrojových systémech, rozsáhlost a proveditelnost procesu, ...).
- **Určení granularity.** V dalším kroku je nutné stanovit granularitu – úroveň podrobnosti – ukazatelů v tabulce faktů pro vybraný proces, což znamená, s jak detailními daty chce uživatel manipulovat (např. záznamy na úrovni dnů, týdnů, měsíců, let – nízká až vysoká úroveň detailu).
- **Identifikace dimenzí.** K vytýčené granularitě je dále potřebné vymezit jednotlivé entity (např. zaměstnanec, produkt,...), které budou sloužit jako popis zkoumaného faktu. Entity musí být s procesem kompatibilní. Ke každé entitě jsou přiřazeny její atributy, jež konkretizují popis faktu (u zaměstnance by se mohlo jednat o jeho jméno, příjmení, pozici, atd.)
- **Identifikace faktů.** K procesu je nutné vyspecifikovat metriky a klíčové ukazatele měřitelnosti (Key Performance Indicators – KPI), jež s ním úzce souvisí. Ty měří výkonnost procesu a jsou pro uživatele důležitým zdrojem svých analýz. Musí být zohledněna agregace faktů vzhledem ke zvolené granularitě. Ve stejné tabulce faktů smí být pouze takové metriky, které mají odpovídající granularitu. V opačném případě musí být v oddělených tabulkách faktů.

Na obrázku 2.4 je možné vidět životní cyklus dimenzionálního modelování, kde Ballard (2006) ještě znázornil verifikaci modelu, fyzický návrh a správu meta dat.

- **Verifikace modelu.** Před samotným fyzickým návrhem musí proběhnout verifikace dimenzionálního modelu, kterou by měly být ověřeny požadavky uživatelů. Při zjištění nesouladu musí proběhnout korekce, tak aby návrh odpovídal zmíněným požadavkům.

- **Fyzický návrh.** Po návrhu a verifikaci následuje zaměření se na další aspekty dimenzionálního modelování. Může jít o ladění výkonu, umístění dat, indexování, partitioning a agregaci dat.

Nad všemi těmito aktivitami jsou spravována meta data. Meta data jsou *data o datech* a jsou uložena ve speciálních úložištích. Ke správnému použití dat je důležité znát jejich strukturu a význam ve zdrojích dat, k tomu jsou určena právě meta data.



Obr. 2.4 Životní cyklus dimenzionálního modelování (Ballard, 2006)

### 2.3.3.Specifikace podnikového procesu a požadavků

V této fázi životního cyklu dimenzionálního modelování je vyhledáván proces včetně požadavků uživatelů, pro který bude navrhován dimenzionální model. Procesem se rozumí opakující se sada činností. Podle Ballarda (2006) musí být brán zřetel na to, že k jednomu procesu může být vyžadováno více dimenzionálních modelů.

Na začátku této činnosti musí být vyspecifikovány a prostudovány veškeré podnikové procesy organizace a sepsány do tabulky procesů. V tabulce jsou jednotlivé procesy ohodnoceny hodnotícími faktory, a to:

- složitost zdrojových systémů,
- dostupnost dat z těchto systémů,
- kvalita dat,

- strategický význam pro podnik.

Firmou je, po dobrém promyšlení, ke každému procesu a faktoru přiřazena hodnota důležitosti z předem stanového intervalu (např. nízká, střední, vysoká). Toto bodování slouží pro lepší prioritizaci procesů, neboť v dalším kroku je pro zpracování dimenzionálního modelu vybrán proces s největším součtem bodů. Postupně jsou vybírány další procesy dle priority (součet hodnotících bodů).

Pomocí bus matrix, která byla vysvětlena v kapitole *Konceptuální datový model*, jsou zachyceny společné dimenze k jednotlivým procesům. Společné dimenze jsou označovány tzv. jako sjednocené dimenze. V rámci komplexního dimenzionálního modelu organizace budou sloužit jako spojovací tabulky těchto procesů. Vyhledávána jsou rovněž sjednocená fakta.

Ke každému procesu musí být rovněž sepsány odpovídající zdrojové systémy včetně jejich detailu, což může být např. název, vlastník, platforma, operační systém a jeho verze, kapacita. Příkladem zdrojového systému může být celopodnikový datový sklad, OLTP systém, nezávislá datová tržiště, atd.

Jelikož je získávání a vybírání správných požadavků složitý ale důležitý proces, je snahou, aby tento proces proběhl optimálně. Podle Ballarda (2006) existuje několik metod, jak toho dosáhnout.

- **5W – 1H**

Jde o sadu otázek, kterými je postupně doptáváno a v angličtině začínají „W“ a „H“ (when, where, who, what, why a how). Např. může jít o otázky typu: „Kdo jsou zainteresovaní lidé, skupiny a organizace?, Jaké funkce potřebují být analyzovány?, Proč jsou data požadována?, Jak je měřena výkonnost procesů?, Jaké faktory určují úspěch nebo neúspěch?“, atd.

- **Zdrojově řízený sběr požadavků**

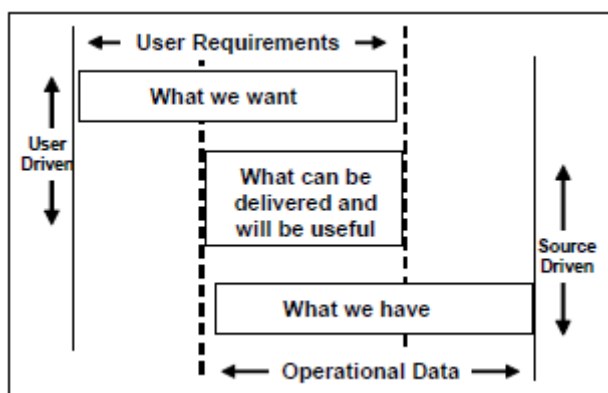
K této metodě definování požadavků je použito zdrojových dat z produkčních provozních systémů. Je založena na analyzování E/R modelu zdrojových dat nebo na aktuálních fyzických záznamech a výběru datových prvků.

- **Uživatelsky řízený sběr požadavků**

K této metodě je využito uživatelů v organizaci, je zjišťována práce a funkce jednotlivých uživatelů. Tento průzkum může být prováděn na různých schůzkách či pohovorech. Dle Kaluži

(2011) mohou být požadavky identifikovány a analyzovány rozhovorem projektanta s uživateli systému, rozbořem písemných materiálů, dotazníky, nebo pozorováním.

Na Obr. 2.5 je možné vidět, jak jsou shromažďovány informace dle zdrojů, nebo uživatelů. Ke zjištění co má být doručeno a co bude užitečné, je potřebné na straně uživatele znát „Co chceme“ a na straně zdrojů „Co máme“.



Obr. 2.5 Zdrojově a uživatelsky řízené shromažďování požadavků (Ballard, 2006)

Během fáze shromažďování požadavků a obchodních potřeb uživatelů jsou tyto údaje dokumentovány a analyzovány. Musí být stanoveno, co bude monitorováno a jaká budou hodnotící kritéria. Metody sběru byly zmíněny v předešlém odstavci. Jednotlivé zjištěné záznamy jsou ve formě otázek a je jim přiřazena významnost (nízká, střední, vysoká, kritická). Tato fáze je důležitá z hlediska pochopení samotného problému, který má být dimenzionálním modelem řešen.

Na konci identifikace požadavků jsou k jednotlivým údajům přiřazeny nejvyšší entity a metriky. Např. ke zjištění průměrného měsíčního prodeje produktu je zapotřebí entita *Měsíc* a *Produkt* a metrika *Množství prodané jednotky*. Zároveň je u entit zaznamenána hierarchie posloupnosti (např. *Produkt* je možné zkoumat z hlediska *Kategorie* nebo *Značky*).

### 2.3.4.Vymezení granularity

Vymezení granularity v tabulce faktů je stěžejní krok v dimenzionálním modelování, neboť granularitou je definováno, co každý řádek v tabulce faktů znamená (jaká je jeho úroveň podrobnosti). Granularita přímo souvisí s úrovní podrobností v dílčích dimenzích – jde o vazbu mezi dimenzemi a tabulkou faktů (např. jsou-li v dimenzi *Čas* formulovány měsíce a v dimenzi *Zaměstnanec* jednotliví zaměstnanci, pak jsou v tabulce faktů vidět měsíční hodnoty za

zaměstnance). Granularita musí být určena ještě před vymezením dimenzí a faktů, protože všechny dimenze a fakta musí odpovídat zvolené granularitě, tzn., že každý řádek musí mít stejný typ dat.

Rozlišuje se mezi **nízkou** úrovní granularity, kdy je tabulka faktu hodně detailní, a **vysokou** úrovní granularity, kdy tabulka faktu obsahuje méně detailní záznamy. (Ballard, 2006) U nízké úrovně granularity většinou není vyžadována žádná agregace, ale na druhou stranu je potřeba více úložného prostoru pro zachování takového množství dat. S vysokou úrovní granularity je právě spojena agregace dat, což má u analytických operací pozitivní dopad. Např. pokud firmy potřebují vidět měsíční tržby a data mají uložena zmíněnou první možností, výpočet je prováděn v reálném čase, ale odezva je oproti druhé variantě příliš pomalá.

Jak již bylo řečeno o přímé souvislosti dimenzí a granularity tabulky faktů, je k pochopení samotné granularity nutné chápat atributy dimenzí. Ballard (2006) tvrdí, že díky vztahu atributů předeek-potomek, je možná sumarizace na několika úrovních – granularita je označována jako nejnížší úroveň sumarizace nebo nejvyšší úroveň detailu.

Pour, Maryška, Novotný (2012) uvádějí doporučení při řešení úrovně podrobnosti:

- ukládat data na nejvyšší úrovni detailu, pokud je to technickým vybavením firem možné,
- data z různých zdrojů zapisovaných do datového skladu převést na stejnou (podobnou) úroveň granularity.

Samotná granularita je výsledkem předem daných potřeb. Již během sběru požadavků by se měly shromažďovat takové dokumenty, které mohou být nápomocné ke stanovení granularity a v následujících krocích dimenzionálního modelování také k dimenzím a faktům. V každé organizaci může jít o rozdílné dokumenty, ale mělo by jít o dokumenty, kde jsou zachycena transakční data, a dokumenty, které přímo i nepřímo souvisí s vytýčeným procesem (např. může jít o faktury, objednávky, skladní listiny, zaměstnanecké karty, ...). Dokumenty mohou být nositelem zásadních informací o zákaznících, produktech, dodavatelích, zaměstnancích, atd., a to i na nejnižších úrovních detailu.

Jelikož je možné, že firma potřebuje k danému procesu v tabulkách faktů sledovat více různých granularit, nebo že u některých sledovaných údajů nemůže být stejná granularita, je potřebné rozdílné granularity od sebe oddělit a vytvořit samostatné tabulky faktů. Zde je ovšem

nutné brát ohled na množství dat, jejich různé agregace, úložný prostor, požadavky na výkon a další technické omezení.

Po vymezení granularity následuje její podrobná kontrola, aby se minimalizoval výskyt chyb v dalších aktivitách životního cyklu a případné předělání původního modelu. Měla by být zajištěna co nejvyšší úroveň detailu, aby podniku mohly být poskytovány detailnější informace (pokud je to možné). K ověření granularitě tabulky faktů jsou předběžně určeny dimenze a fakta, jež budou zvolenou granularitu popisovat.

### 2.3.5. Identifikace dimenzí

Cílem identifikace dimenzí v životním cyklu dimenzionálního modelování je nalezení všech potřebných dimenzí, které mají odpovídající granularitu, jež byla v předešlém kroku zvolena. Díky vytvoření závěrečné zprávy v kapitole *Vymezení granularity*, kde jsou zaznamenány veškeré granularity podnikových procesů v tabulce faktů a předběžným stanovením dimenzí, je v tomto kroku mnohem jednodušší nalézt všechny dimenze.

Dimenze představují jednotlivé objekty reálného světa a společně s atributy slouží jako textový či numerický popis záznamů v tabulce faktů. Dimenze ve většině případů obsahují data na nejnižší úrovni detailu – granularita dimenzí. Použití atributů dimenzí může mít dvojí smysl:

- popisná informace metriky (kdo, co, kde, kdy, jak a proč),
- jak sumarizovat data v tabulce faktů, aby analytikům byla poskytnuta užitečná informace.

#### Atributy dimenzí

Dimenze jsou charakteristické tím, že mohou obsahovat mnoho atributů. Počet atributů se může lišit a dimenze jich mohou nabývat padesát i víc. Na druhou stranu nemají tolik řádků. Těmito charakteristikami jsou dimenze rozdílné od tabulky faktů, která má naopak obrovské množství řádků a menší počet atributů.

Atributy jsou považovány za hlavní zdroj seskupování, filtrování, omezování záznamů, a to jak v dotazech, tak v sestavách reportů. Rovněž slouží jako popisky v reportech, atd., proto jim musí být věnována velká pozornost. Pojmenování atributů může vypadat jako primitivní aktivita, ovšem opak je pravdou, neboť atributy jsou úzce spjaty s kvalitou datového skladu jako celku

a odvíjí se od nich celý jeho budoucí chod. Dle Kimballa a Rosse (2013) představují atributy dimenzí hlavní roli pro DW/BI <sup>3</sup> systémy zejména proto, že:

- jsou zdrojem všech omezení a popisů reportů,
- jsou kritické pro chápání a použitelnost DW/BI systémů.

Aby byly dodrženy předešlé body měly by se názvy popisných atributů (výjimka primárních klíčů dimenzí) skládat z reálných slov. Obsah atributu by měl být v názvu jasně a stručně vystižen bez použití zkratk či kódů. Je to proto, že s reporty pracují uživatelé, kteří by se ve zkratkách a kódech nevyznali. Mnohdy se však kódům nelze vyhnout. K porozumění kódů by měla být vytvořena zvláštní dimenze, jež bude zdrojem pro standardní dekódování. Dojde tak ke zpřehlednění a k usnadnění použitelnosti zakódovaných atributů.

Tabulka faktů může být popsána i numerickými atributy dimenzí (např. cena produktu). Někdy je obtížné rozhodnout, zda má být numerický atribut zařazen do dimenze či do tabulky faktů. Determinujícím faktorem je proměnlivost (vznikající změně) daného atributu. Je-li změna častá, měl by být atribut součástí tabulky faktů. Nevzniká-li tolik změn, jejich rozptýl je více či méně konstantní a jsou používány k omezujícím podmínkám dotazů, pak by měl být tento atribut zahrnut do dimenze.

### **Klíče dimenzí (primární / cizí / umělé)**

V dimenzionálním modelování by se dle Ballard (2006) měly v dimenzích používat tzv. umělé klíče (*surrogate key*), které mají zajišťovat stabilitu primárních klíčů. Jsou to takové klíče, které plní funkci primárních klíčů v datovém skladu, namísto primárních klíčů, jež jsou v původních systémech. Každá dimenze by měla disponovat alespoň jedním umělým (primárním) klíčem, kterým jsou propojeny s odpovídající tabulkou faktů, příp. jinou dimenzí. Primární klíče v tabulkách dimenzí jsou v tabulkách faktů cizí klíče. Jednotlivé dimenze jsou k tabulce faktů ve vztahu 1:N. Umělé klíče se využívají k identifikaci instance nebo entity uvnitř tabulky dimenze. Každý neklíčový atribut by měl být pouze v jedné dimenzionální tabulce.

Existuje mnoho důvodů, proč by v dimenzích měly být umělé klíče:

- do datového skladu nejsou propisovány změny umělých klíčů v původních databázích,
- efektivnější a kvalitnější konsolidace u překrývajících se klíčů transakčních systémů,

---

<sup>3</sup> Systémy datového skladu a Business Intelligence



- zajištění lepšího výkonu datového skladu způsobenou menším rozsahem atributů,
- úspora v podobě složených klíčů v tabulce faktů,
- jsou nápomocné u neočekávaných situací (např. *není známo*) a mnoho dalších.

Na jednu tabulku faktů by optimálně mělo být napojeno patnáct dimenzí. Pokud tomu tak není, je vhodné spojit několik dimenzí do jedné. Sloučit se mohou pouze dimenze se souvisejícími atributy. (Pour, Maryška, Novotný, 2012) Často je diskutováno, zda má být vytvořeno méně komplexních dimenzí, nebo více jednodušších. Menší počet dimenzí je jednodušší pro pochopení uživatelů, pro správu a pro provádění úprav.

### **Dimenze času**

Nezávisle vystupující a nejčastěji používanou dimenzí je dimenze času, která má apriori danou hierarchii (rok, měsíc, čtvrtletí, týden, den, ...). Záleží již na samotné firmě, jak dimenzi času použije a jakou úroveň detailu využije. V dimenzionálním modelu se téměř vždy vyskytuje dimenze času, neboť organizace chce sledovat metriky za nějaké časové období. Data do dimenze času jsou čerpána buď z kalendáře podniku, nebo je kalendář vytvořen manuálně na období, jímž budou pokryty sledované hodnoty.

Dimenze data a dimenze času by měly být od sebe odděleny a v dimenzionálním modelu by měly tvořit samostatné tabulky. U spojených tabulek by při rozkladu času docházelo ke zbytečnému opakování data.

### **Speciální dimenze**

Podle různých situací je se možné setkat s několika možnými typy dimenzí např.:

- **Sdílené dimenze** (*Conformed Dimension*) – dvě dimenze jsou označeny za shodné, pokud jsou jejich názvy atributů stejné a mají stejný význam. Jsou určeny k propojení datových tržišť.
- **„Parent-child“ dimenze** – jde o dimenze, které jsou ve vztahu předeek a potomek. Např. může jít o zachycení vztahu nadřízenosti a podřízenosti.
- **Dimenze různých rolí** (*Role-playing Dimension*) – jsou to dimenze svým obsahem stejné, ale vzhledem k popisu tabulky faktů mají poskytovat jiný obsah. Např. různý kontext data.
- **Sběrná dimenze** (*Junk Dimension*) – jde o tabulku s relativně malým počtem prvků, které byly vyjmuty z tabulky faktů. Může jít o kombinace různých příznaků způsobující v tabulce faktů nárůst záznamů.

- **Pomalů mēnící se dimenze** (*Slowly Changing Dimensions, SCD*). SCD je řešeno zachycování změn v dimenzích (přidávání, rušení prvků, změny názvů prvků nebo změna zařazení). Pro SCD existují tři vzory, jak se změnami zacházet. *Typ 1* – je představován přepsáním původního řádku danou změnou. *Typ 2* – změna je zachycena přidáním nového řádku do dimenzionální tabulky. *Typ 3* – změny jsou zaznamenávány do zvláštního sloupce, kde v jednom sloupci je původní hodnota a v druhém aktuální.
- **Rychle mēnící se dimenze** (*Rapidly Changing Dimensions, RCD*). Jde o dimenze, u kterých je možná změna na denní, týdenní či měsíční bázi. Zde jednotlivé typy SCD nejsou přijatelné (obrovské množství záznamů). K původní dimenzi je ve vztahu 1:N vytvořena tzv. mini-dimenze, v níž je zachycována podmnožina často mēnících se atributů.

### 2.3.6. Identifikace faktů

Cílem identifikace faktů je nalezení všech metrik a indikátorů, jež mají být v daném podnikovém procesu zkoumány. Ke snadnějšímu nalezení všech faktů je určena již dříve a předběžně zpracovaná tabulka (zpráva) všech granularit a jejich možných dimenzí a faktů. K zjištění případných dalších faktů jsou použity E/R modely. Do tabulky faktů mohou spadat pouze fakta se stejnou granularitou (ostatní musí být odděleny).

Sloupce tabulky faktů jsou reprezentovány cizími klíči a hodnotami vypočtených metrik. Cizí klíče dimenzí jsou vazbou na dimenze popisující vybraná fakta. Ve sloupci tabulky faktů představují hodnoty cizích klíčů konkrétní prvky připojených dimenzí. Primární klíč tabulky faktů je právě složen z těchto cizích klíčů, tzv. *složený primární klíč*. Tabulka faktů je k jednotlivým dimenzím ve vztahu M:N. Řádky tabulky faktů reprezentují jednotlivá měření vzhledem ke zvoleným dimenzím (např. hodnota za zemi, kraj, město, prodejnu, zaměstnance). Záznamy jsou na nejnižší úrovni detailu. Datový sklad je z devadesáti procent tvořen právě hodnotami z tabulek faktů, zbylých deset procent je zastoupeno hodnotami dimenzí.

Do tabulky faktů by neměla být doplňována samostatná hodnota, jež by měla plnit funkci primárního klíče (primární klíč je tvořen cizími klíči dimenzí) – výjimka je pouze v případě, že složený primární klíč by nebyl jednoznačný. Rovněž by cizí klíče neměly mít hodnotu NULL (nemá vypovídací schopnost pro uživatele), tato hodnota by měla být nahrazena „N“, příp. nulou. Hodnota NULL se může vyskytovat ve sloupcích, které jsou použity do výpočtu, příp. je nad daným sloupcem prováděna agregace.

Hodnoty v tabulce faktů (vyjma cizích klíčů) jsou tvořeny buď hodnotami z původních databází (např. cena produktu, nachází-li se v tabulce faktů) nebo jsou před uložením do datového skladu (datového tržiště) předkalkulovány (např. čistá tržba). Potřebné hodnoty mohou být ale také vypočteny až v analytických nástrojích.

Kimball a Ross (2013) dělí tabulky faktů do tří kategorií dle dané úrovně granularity:

1. **Transakční** – tento typ tabulky faktů, je úzce spojen s produkčními systémy a ukládá záznamy na nejvyšší úrovni detailu. Údaje jsou pouze vkládány.
2. **Periodické údaje** – v tomto typu tabulky jsou uchovávány hodnoty metrik za pevně stanovené období. Obsahuje méně záznamů než transakční tabulka. Údaje jsou rovněž jenom vkládány.
3. **Kumulativní údaje** – data v tomto druhu tabulky nemusí být v pevně ohraničeném časovém intervalu, příp. mohou být zaznamenávány aktualizované údaje sledovaného procesu. Hodnoty mohou být jak vkládány, tak aktualizovány.

Fakta se vzhledem k **agregaci** dat dělí podle Ballarda (2006) na:

1. **aditivní** – agregace tohoto typu je vždy smysluplná (např. suma tržeb za měsíc),
2. **semi-aditivní** – agregace dává smysl pouze u určitých dimenzí (např. suma prodaných produktů právě za daný produkt, smysl již nedává suma prodaných produktů za měsíc),
3. **ne-aditivní** – agregace nemá žádný smysl (např. suma za procentní slevu).

V dimenzionálním modelu se mohou vyskytovat i tabulky faktů, které nemají žádné ukazatele, jde o tzv. bezfaktovou tabulku faktů (*factless fact table*). Tyto tabulky mají zachycovat počet uskutečněných událostí.

### 2.3.7.Fyzický návrh dimenzionálního modelu

V této podkapitole jsou popsány pojmy jako agregace, indexace a návrh rozdělení tabulek (*partitioning*). Jde o část kroků týkající se fyzického návrhu dimenzionálního modelování.

K fyzickému návrhu je použito již dříve vytvořených modelů na konceptuální a logické úrovni popsaných v kapitole 2.2.2 a 2.2.3. Samotná realizace je uskutečněna již v konkrétním, vybraném databázovém prostředí. Jsou vytvářeny fyzické tabulky (identifikované dimenze a tabulky faktů, o kterých bylo diskutováno v kapitole 2.3.5 a 2.3.6) s danými atributy a různými charakteristikami pro zajištění výkonnosti datového úložiště. Důraz je kladen na uložení dat

a minimalizování vstupně-výstupních operací, které se výrazně podílí na době odezvy jednotlivých požadavků. K poskytnutí co nejrychlejšího řešení je využito lepší organizace, uspořádání a označení dat. Novotný, Pour, Slánský (2005) uvádí např. *verifikace, případné úpravy či nový návrh granularity dat, návrh rozdělení tabulek, slučování tabulek, řešení datových polí, dělení dat podle pravděpodobného přístupu, agregace, indexace*.

- **Verifikace, případné úpravy či nový návrh granularity dat** – o granularitě již bylo hovořeno v kapitole 2.3.4 a byl zmíněn i vliv granularity na výkon datového úložiště.
- **Návrh rozdělení tabulek** (*partitioning*) – jak už je z názvu patrné, jde o rozdělení rozsáhlé tabulky do více malých - lépe zvládnutelných tabulek. Oddělené tabulky mají vliv na lepší výkon právě proto, že jsou lépe manipulovatelné. V případě rozdělených tabulek se při použití dotazů již nemusí procházet celá rozsáhlá tabulka, ale pouze některé její části odpovídající podmínce. Je možné využít horizontální a vertikální dělení tabulek, kde horizontální je rozdělení podle řádků a vertikální podle sloupců.
- **Slučováním** (*merging*) **tabulek** – cílem je snížení odezvy na položený dotaz, neboť dotaz nemusí procházet tolika tabulkami. Slučováním tabulek dochází k denormalizaci dat, která v případě dimenzionálního modelu nejsou problémem jako u relačních databází.
- **Řešení datových polí** (*arrays of data*). U sekvenčně uspořádaných dat je možné použít ukládání jednotlivých dat do řádků tak, jak na sebe logicky navazují. Dochází k urychlení vstupně-výstupních operací.
- **Dělení dat podle pravděpodobných přístupů** – dlouhé záznamy s různorodými položkami jsou podle pravděpodobnosti přístupu rozděleny do různých částí. U položených dotazů se pak přistupuje pouze k jednotlivým částem, což snižuje dobu přenášení záznamů.
- **Agregace** – jde o různé předkalkulační přípravy v dimenzionálním modelu např. v podobě součtu, průměru, počtu, ..., za určitou skupinu dat (např. součet prodaných produktů za měsíc). Agregáty jsou silným nástrojem pro zrychlení dotazovacích procesů, ovšem agregace je také spojena s náklady na úložný prostor, na vytvoření a údržbu ETL procesu spojeným s agregačními tabulkami. Tím, že se dimenze skládají z atributů (mohou vznikat hierarchie), je možné sumarizace provádět na různých úrovních, což umožňuje procházet data v úrovních nahoru a dolů (*drill down, drill up*), podle toho, jak uživatel potřebuje.
- **Indexování záznamů** v tabulkách dimenzionálního modelu je určeno k uchování informací o tom, kde se daný záznam fyzicky nachází. Po zadání dotazu není prohledávána celá

tabulka, ale je načten pouze index záznamů a z tabulky jsou tyto záznamy získány. Indexy je mnohonásobně urychlena doba odezvy při pokládání dotazů. Ballard (2006) dělí indexy na unikátní, neunikátní, indexy B-Stromu a bitmapové indexy.

- **Unikátní** – každý řádek je vůči ostatním hodnotám indexu unikátní (vyskytuje se tam pouze jednou).
- **Neunikátní** – hodnota indexu může odkazovat na jeden nebo více řádků v tabulce.

Dle fyzické organizace se indexy dělí na:

- **B-Strom** – indexy se v této struktuře nacházejí na konečných uzlech (listech), v nichž je rovněž uložen ukazatel na fyzický záznam. Je používán v modelech s vysokým stupněm kardinality.
- **Bitmap** – index tohoto typu je uložen v každém bitu, který slouží jako ukazatel na daný řádek tabulky.

#### **Clusterovaný index**

Indexy B-Stromu mohou být clusterované, čímž je myšleno, že fyzické řádky tabulky jsou řazeny podle indexu. V případě čtení tabulek podle indexu, dochází ke snížení vstupně-výstupních operací. Clusterovaný index je nejučinnější, jestliže po jeho vytvoření už nenastanou v tabulce žádné změny, jako např. přidání nového řádku, úprava nebo smazání původního řádku. Tyto operace jsou příčinou snížení efektivnosti clusterovaného indexu, protože řádky nejsou příliš dlouho ve fyzickém pořadí indexu. K udržení efektivnosti clusterovaného indexu by indexy měly být vytvořeny znovu, což může být u rozsáhlých tabulek nákladné. Proto je možné indexy vložit do oddělených logických míst.

### **2.3.8. Shrnutí**

Kapitola *Dimenzionální modelování* je stěžejní kapitolou diplomové práce, v níž je mimo jiné zmíněna metodologie R. Kimballa životní cyklus vývoje dimenzionálního modelu. V praktické části práce půjde o vytvoření dimenzionálního modelu podle popsaného životního cyklu. V první řadě budou vytýčeny procesy, pro které bude dimenzionální model navrhován, včetně požadavků uživatelů. V druhém kroku bude stanovena granularita tabulky faktů a sepsána do výstupní tabulky spolu s předběžnými dimenzemi a tabulkami faktů. V třetí fázi budou identifikovány všechny dimenze s jejich atributy a umělými klíči. Bude-li zapotřebí, budou vymezeny speciální dimenze.

Následně bude specifikována tabulka faktů včetně zkoumaných metrik. V poslední řadě bude u vybraných atributů zavedena agregace, indexování a další možné operace fyzického návrhu.

Všechny zjištěné informace budou zobrazeny pomocí sémantického, konceptuálního a logického modelu, jak již bylo řečeno v kapitole *Datové modelování*.

## **2.4. Reprezentace a uložení dat multidimenzionálního modelu**

Jelikož je multidimenzionální model odlišný od klasických databází tím, že je využito dimenzí a tabulek faktů k různým pohledům na data, je potřeba použít jiné reprezentace a uložení dat. Multidimenzionální model může být uložen a reprezentován ve formě relační databáze, nebo uložen ve formě multidimenzionální databáze (Online Analytical Processing - OLAP) a reprezentován ve formě fyzické datové kostky.

### **2.4.1. Relační databáze**

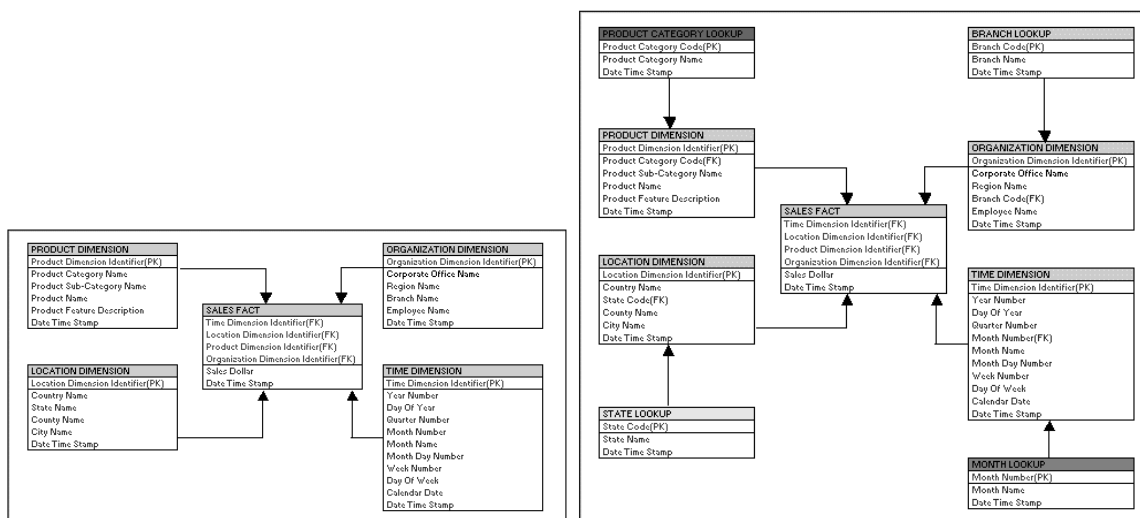
Multidimenzionální model může být z hlediska propojení a vztahů mezi dimenzemi a tabulkami faktů navrhnut jako schéma hvězdy (*STAR schema*), schéma vločky (*SNOWFLAKE schema*), nebo kombinací těchto dvou schémat jako schéma galaxie (*GALAXY schema*).

#### **Schéma hvězdy**

Jak je možné vidět na Obr. 2.6a, dimenzionální model skutečně vypadá jako hvězda. Tabulka faktů je středem modelu a dimenzionální tabulky jsou s tabulkou faktů ve vztahu ve tvaru kruhu. Dimenzionální tabulky jsou v nenormalizované formě, s čímž je spojeno rychlejší vyhledávání v těchto tabulkách, neboť v tomto konceptu není tolik spojení (*join*). Dochází tudíž k rychlejšímu zpracování, ale na úkor úrovně detailnosti, která nemůže být tak podrobná.

#### **Schéma vločky**

Schéma vločky je typické tím, že do dimenzionálního modelu je zahrnuta částečná nebo úplná normalizace. Ta je vytvořena tak, že na jednotlivé dimenze jsou napojeny dimenze další, které jsou použity pro rozšíření původní dimenze spojenou s tabulkou faktů. Model je přehlednější a umožňuje snadnější správu dimenzí, ale při volání dotazů musí projít všechny tabulky a doba odezvy se zvyšuje. Jak vypadá příklad pro schéma vločky, je možné vidět na Obr. 2.6b.



Obr. 2.6a Schéma hvězdy (LEARNDWBI, 2015a) 2.6 b Schéma vločky (LEARNDWBI, 2015b)

### Schéma galaxie

Schéma galaxie je představováno kombinací schémat hvězdy a vločky. Toto schéma se následně skládá z několika dimenzí a tabulek faktů. Různé tabulky faktů mohou mít na sebe napojeny stejné dimenze, tzv. sdílené dimenze (*conformed dimension*). Jak již bylo zmíněno v kapitole 2.3.5, jsou pomocí sdílených dimenzí spojována datová tržiště, čímž vzniká celopodnikový datový sklad. Je tím vyjádřen fakt z praxe, kdy na různou informaci může být pohlíženo ze stejného úhlu pohledu.

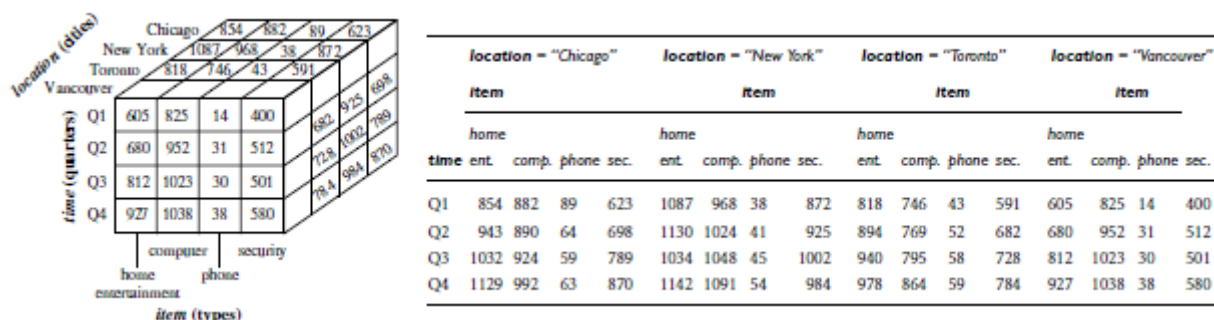
## 2.4.2. Multidimenzionální databáze

Data v multidimenzionální databázi jsou oproti relačním databázím uložena jiným způsobem, a to v podobě *datové kostky*. Multidimenzionální databáze je také zaměřena na konkrétní řešení problém podniku. Fakta v datové kostce jsou již určitým způsobem předkalkulovaná - na předem vyspecifikované požadavky může jít o různou úroveň sumarizace či průměru. Tím je dosaženo rychlejší reakce na položené dotazy, ale jsou zapotřebí větší nároky na kapacitu úložiště.

V datové kostce může být uloženo několik dimenzí – v nejjednodušším případě jde o dvě dimenze a v případě rozsáhlé datové kostky až o  $n$  dimenzí. Na agregovaná data lze také pohlížet z různých oddělených datových kostek, a to v případě, že počet dimenzí je větší než čtyři, nebo že daná fakta již z výše uvedených důvodů nelze agregovat.

U multidimenzionálních databází je nevhodné uchovávat data na nejvyšší úrovni granularity, neboť hlavním účelem dimenzionálního modelu je zkoumat různě agregovaná fakta z odlišných úhlů pohledů. Většinou ani není možné mít uložená data na vysoké úrovni detailu, protože by došlo k zahlcení databáze a efektivnost multidimenzionálních databází by byla velmi nízká a manažerské rozhodování by nemohlo být tak rychlé, jak je očekáváno.

Na Obr. 2.7 je možné vidět příklad multidimenzionální databáze, kde je datová kostka složena ze tří dimenzí, a jednotlivé buňky kostky jsou agregovaná fakta.



Obr. 2.7 Příklad datové kostky a agregovaného výstupu v podobě tabulky (Han a Kamber, 2006)

V Tab. 2.2 jsou shrnuty výhody a nevýhody relačního a multidimenzionálního modelu.

Model	Výhody	Nevýhody
<b>Relační model</b>	<ul style="list-style-type: none"> <li>• zvyklost pracovníků na relační model</li> <li>• dostupný SW a vývojové nástroje pro vývoj a ladění aplikací a pro generování reportů</li> <li>• použitelnost v OLTP i DW</li> </ul>	<ul style="list-style-type: none"> <li>• absence komplexních analytických nástrojů</li> <li>• omezení objemu údajů, ke kterým lze v čase rozumně přistupovat</li> </ul>
Model	Výhody	Nevýhody
<b>Multi-dimenz. model</b>	<ul style="list-style-type: none"> <li>• rychlý komplexní přístup k velkému objemu údajů</li> <li>• přístup k multidimenzionálním a relačním datovým strukturám</li> <li>• možnost komplexních analýz</li> <li>• silné schopnosti pro modelování a prognózy</li> </ul>	<ul style="list-style-type: none"> <li>• problémy při změně dimenzí</li> <li>• vyšší nároky na kapacitu úložiště</li> </ul>

Tab. 2.2 Výhody a nevýhody relačního a multidimenzionálního modelu (Lacko, 2003)



### 2.4.3. Shrnutí

Cílem kapitoly *Reprezentace a uložení dat multidimenzionálního modelu* bylo popsání možností reprezentace a uložení dat včetně jejich charakteristik a nalezení výhod a nevýhod jednotlivých alternativ. Multidimenzionální model, navržený v praktické části diplomové práce, bude reprezentován a uložen pomocí jednoho z těchto řešení.

## 2.5. Komponenty pro analýzu a reporting

Pro analýzu a reporting je možné využít řadu softwarových nástrojů, které mohou nabízet řadu prostředků, jak pro analytické operace, tak pro prezentaci a vizualizaci výsledků, jež se odvíjí od položených dotazů. Ty se mohou lišit dle použitého databázového prostředí např. SQL příkazy. Jednotlivé příkazy mohou probíhat standardně v předem stanovených časových intervalech, ale uživatelem mohou být tvořeny i ad-hoc (náhodné) příkazy pro specifický výstup.

### 2.5.1. Multidimenzionální analýzy

Smyslem multidimenzionálních analýz je procházení různých úrovní detailů v dimenzionálně uspořádaném modelu. Jde o jednu z funkcí tzv. online analytického zpracování, neboli *On-line Analytical Processing (OLAP)*, o kterém již bylo psáno v souvislosti s uložením multidimenzionálních dat. Uživateli je v reportingových nástrojích vždy zobrazena sestava ukazatelů ve výchozím zobrazení (nízký detail) a díky OLAP je schopen podrobnější analýzy těchto ukazatelů.

Při použití OLAP nástrojů je možné nad daty provádět tyto **operace**:

- **Slicing** – jde o řez, který je proveden na datové kostce vzhledem k jednomu atributu dimenze. Poté jsou vyfiltrovány agregované hodnoty pro daný měsíc, příp. jiný atribut.
- **Dicing** – také jde o řez datovou kostkou, ale řez je možno provést přes více atributů dimenze.
- **Drill-down** – pro tuto operaci je typické zvyšování úrovně detailu, kdy se postupuje od nejvíce agregovaných hodnot v tabulce faktů až k méně agregovaným hodnotám. Může jít i o přidání další dimenze, dojde tím opět k větší podrobnosti faktů.
- **Drill-up** – jde o opačnou operaci jako je drill-down. Tabulka faktů je zkoumána od detailnějších údajů k méně detailním, tedy k větší agregaci faktů. Podobně jako v předešlé operaci je zde naopak možnost odebrání některé z dimenzí.

- **Drill-across** – podle Ballarda (2006) jde o metodu, ve které se prochází z jedné dimenze do druhé a předem musí být známa cesta procházení, tzn., jsou zkoumány nejnižší úrovně detailu.
- **Pivoting** – jde o operaci, u níž je možné měnit úhel pohledu na určitou skupinu dat (např. v případě kontingenční tabulky je na úrovni řádků vyměněna dimenze za jinou a u data ve sloupcích budou mít jinou vypovídací schopnost).

### 2.5.2. Analytické a reportingové nástroje

Dotazovací a reportingové nástroje musí být přesně navrženy pro jednotlivé typy manažerů na strategické, taktické a operativní úrovni, neboť analytické a reportingové nástroje přímo souvisí s manažerskými funkcemi, a to plánováním, organizováním, personalistikou, řízením a koordinováním. Rovněž je zapotřebí, aby klíčové informace byly zpřístupněny pouze vybraným manažerům, a aby návrhy reportingových nástrojů byly navrženy tak, aby poskytovaly jasné, jednoduché, přehledné, rychlé a uživatelsky přijatelné „zpravodajství“ v různých formách zobrazení (tabulky, grafy, mapy, atd.). Nástroje by měly být vysoce flexibilní (při změnách dimenzí). Měly by mít sadu standardních reportů (pro zrychlení doby odezvy). Mohou vstupovat do jiných technologických prostředí (relační DW, OLAP). Komunikace s uživateli musí být intuitivní a efektivní. Každý použitý nástroj by měl mít svou dokumentaci.

Mezi standardní uživatelská rozhraní systému BI jsou reporty, dashboardy, skórovací karty, podnikové portály ad.

#### Reporty, sestavy

Reporty a sestavy je možné vytvářet pomocí různých softwarových nástrojů. Mohou být rozesílány vytýčeným osobám pro prezentaci výsledků organizace. Data je možné čerpat z různých systémů (Online Transaction Processing - OLTP, OLAP), kdy výstupem mohou být statické, či polo-statické reporty. Obsah reportů je dán požadavky uživatelů. Reporty mohou mít standardizovanou formu, ale za pomoci různých nástrojů mohou být vytvářeny i ad-hoc reporty (nestandardní).

#### Dashboardy

Dashboardy jsou vícevrstvé systémy pro řízení výkonnosti organizace. Jsou založeny na Business Intelligence a integraci dat z různých infrastruktur, které organizaci umožňují měřit, monitorovat a řídit podnikové aktivity použitím finančních i nefinančních opatření. (Eckerson,

2006) Obsah a funkce dashboardů se liší podle manažerských úrovní – strategická, taktická a operativní úroveň. Více o dashboardech je popsáno v kapitole 2.5.3.

### **Skórovací karty**

Skórovacími kartami (nebo také *Balanced Scorecards* - BSC) jsou zachycovány metriky ve čtyřech klíčových oblastech podniku – zaměstnanci, zákazníci, finanční a operativní ukazatele. Zaznamenávají jsou aktuální a očekávané hodnoty metrik, aby bylo možné provádět srovnání aktuální situace organizace. Většinou jsou skórovací karty využívány na strategické úrovni podniku.

### **Podnikové portály**

Podnikové portály jsou rozšířenou a flexibilnější formou poskytování reportů v podobě komplexního rozhraní (možná nadstavba dashboardů). Pracovníkům organizace jsou tímto poskytovány analytické informace v jednotném prostředí – snadnější přístup k reportům. Integrovaní dat může být rovněž z různých zdrojů. Touto formou mohou být informace sdíleny i uživatelům mimo organizace jako např. zákazníkům. Výstupy však mají odlišný obsah od interních informací. Podle Němce (2014) mohou portály obsahovat interaktivní analytické reporty, statické reporty, grafické indikátory výjimek, odkazy na relevantní aplikace zdroje informací, aplikace zaměřené na specializovanou oblast, činnost nebo funkci, atd.

## **2.5.3. Dashboardy na principu User Experience**

Při navrhování dashboardů, ale i jiných aplikací, nástrojů, s kterými budou ve styku uživatelé, je vhodné, aby navrhování probíhalo na principech *User Experience* (dále UX). Gube (2010) tvrdí, že UX představuje pocit uživatele spojený s prací v daném systému. Jde o to, zaměřit se na potřeby uživatele a uspokojit jeho potřeby při hledání informací. UX má úzkou souvislost s *User Interface* <sup>4</sup>(dále UI) - podmnožina UX. UX není pouze o vzhledu stránky, ale i tom jak jsou informace dostupné, jak rychle je možné, je najít, jaké jsou možnosti na dashboardu, jaká je doba odezvy při práci s dashboardem, jestli uživatel najde to, co hledá, jak je přistupováno ke změnám a mnoho dalších. Cílem je také sbírat zpětnou vazbu od uživatelů a přepracovat dashboardy do vyhovující formy, což je snahou minimalizovat.

---

<sup>4</sup> Uživatelské rozhraní – jde o všechny prostředky, které uživatel využívá ke komunikaci s počítačem (strojem)

K dosažení pozitivního hodnocení je v první fázi návrhu dashboardů důležité vytýčit uživatele jednotlivých procesů (týmy operačních procesů, servis manažery, produkt manažera, atd.), získání a hlavně pochopení uživatelských požadavků, s jakým úmyslem bude dashboard použit a co bude výsledkem použití dashboardů. V podniku potřebují různé skupiny lidí pozorovat odlišné druhy informací.

Následně by měly být všechny sesbírané požadavky a připomínky rozděleny do skupin podle logického uspořádání (souvislost hodnot, souvislost uživatelů, atd.) na dashboardech. Jaká by na nich měla být skupina informací a pro jaké uživatele. Poté je možné navrhovat jednotlivé dashboardy.

Při návrhu dashboardů je důležité, aby uživatele dashboard oslovil, aby uživatel našel, co hledal, a aby porozuměl prezentovanému obsahu (většinou jde o důležité informace). Dashboardy ve většině případů slouží ke sledování KPI, proto by, i při vší možné interaktivitě a OLAP technikám, měly být ve výchozím nastavení zobrazeny právě nastavené ukazatele. (Gibson, 2014) Dashboardy by měly být navrženy co nejjednodušší a nejjasnější formou, neměly by být „přeplácené“ a zahlcené nesmyslnými funkcemi. Jde o to, že co je jednoduché, použitelné a uživatelsky přívětivé je mnohdy více efektivní. Prezentované hodnoty by na stránce měly být rozděleny do jednotlivých karet. Navigace na dashboardu by měla být intuitivní (navrácení zpět, na domovskou stránku, atd.)

K prezentaci ukazatelů, výsledků, trendů by mělo být použito vhodných grafických či tabulkových prostředků jako grafy, budíky, časové osy, tabulky (v tabulkách by mělo být pouze shrnutí základních informací, ne rozsáhlé kontingenční tabulky). Uživateli by mělo být umožněno sledování změn (např. změna oproti minulému měsíci vzhledem k pozorovanému období), aby mohl provádět rozhodnutí na základě porovnání. Měly by být optimálně použity funkce pro multidimenzionální analýzy (pro sledování detailnějších informací). Zásadní je postupovat od shora dolů. (Tobolka, 2013) Ačkoli se to nemusí zdát jako klíčový bod, je vhodné pro upozorňování a informování uživatele využít barvy a značky (např. semaforey pro splnění / nesplnění cíle či jiné barevné odlišení odchylky). Varování může být poskytnuté i jinou formou, manažerovi může být zaslán e-mail nebo SMS se zprávou o neočekávané změně (pro rychlé rozhodnutí manažera). Ke každému prvku dashboardu, by měl být někde uložen krátký popis, jaká informace je daným prvkem poskytována nebo jak byl ukazatel vypočten (např. ve vysouvací ikoně).

Data, jež jsou čerpána do dashboardů, by měla být např. nahrávána do datového skladu během noci, příp. real-time, aby uživatelé měli včas aktuální informace. Dashboardy by měly být přístupné z různých zařízení (počítače, tablety, mobilní telefony, ...).

#### **2.5.4. Shrnutí**

Komponenty pro analýzu a reporting jsou podstatnou částí vzhledem ke koncovým uživatelům (např. dashboardy). Je to „gró“ celého dimenzionálního modelu. Manažeri potřebují sledovat čísla, křivky, barvičky,..., aby mohli rozhodovat o budoucím vývoji podniku. Ke své práci potřebují ve správný čas, na správném místě, správná čísla, možnosti hlubšího zkoumání, pěkný a přehledný design dle jejich požadavků. Nezajímá je, co je na pozadí dashboardů.

Informace této kapitoly budou v praktické části diplomové práce uplatněny ve formě navržených dashboardů pro jednotlivé role manažerů včetně různých možností operací nad daty a designu na principu User Experience.

### **3. Analýza současného stavu zpracování a poskytování reportů v oddělení IT společnosti**

V kapitole týkající se analýzy současného stavu zpracování a poskytování reportů v oddělení IT společnosti je popsána samotná firma, zainteresované oddělení a procesy v něm. Poslední část je věnována vymezení rolí manažerů poskytovaných reportů, jaké reporty, metriky a KPI jsou zpracovávány. V neposlední řadě jaké jsou nyní zapotřebí nástroje, procesy, zdroje a úložiště.

#### **3.1. Popis firmy a firemního prostředí**

Tato kapitola je zaměřena na popis podniku jako celku – základní charakteristika firmy, organizační struktura a procesy v podniku.

##### **3.1.1. Základní charakteristika firmy**

Společnost, pro kterou je zpracovávána tato diplomová práce, poskytuje služby<sup>5</sup> v oblasti informačních technologií (dále IT). Firma byla založena v roce 1968 a její působnost je ve více než dvaceti zemích. V České republice je od roku 2001 se sídly v Praze, Brně a Ostravě. V Ostravě je firmou zaměstnáno okolo dvou tisíc zaměstnanců, celkově ve firmě pracuje patnáct tisíc expertů. Snahou firmy je zaměřovat se na takové trhy, na kterých je šance obdržet alespoň třetí místo v žebříčku IT firem.

Firmou je zajišťována řada služeb - aby v různých zemích celého světa fungovaly firemní procesy - internetové obchody, banky, veřejná doprava, vládní úřady, mohly být prováděny obchodní transakce, mohl být vyráběn papír ad (procesy kontroluje a řídí). Organizace do svých řešení kombinuje nejlepší technologie, které mají plně uspokojit specifické potřeby klientů. Hlavní náplní je vytvářet a zvyšovat obchodní hodnotu pro zákazníka. Od většiny konkurentů, kteří jsou poskytovateli jedné služby, je tato firma odlišná především tím, že poskytuje komplexní IT služby. Zákazníkům jsou nabízeny služby s kompletním životním cyklem – od konzultace, integrace, spuštění, údržbu až po modernizaci systémů. Firmou jsou poskytovány služby pro vývoj a správu aplikací, aplikační outsourcing, platformy obchodních aplikací, systémy pro řízení procesů,

---

<sup>5</sup> Dle Cartlidge a spol (2007) je služba prostředek dodávání hodnoty zákazníkovi tím, že zprostředkovává výstupy, jichž chce zákazník dosáhnout, aniž by vlastnil specifické náklady a rizika.

konzultační služby, portálová řešení, Business Intelligence, služby zabývající se infrastrukturou a další. Těchto služeb je využíváno širokou škálou průmyslových odvětví. Jde o energetické společnosti, finanční služby, dřevařský průmysl, zdravotnické a sociální služby, logistiku, výrobní podniky, mediální společnosti, ropný a plynárenský průmysl, veřejný sektor, maloobchody, velkoobchody a telekomunikační průmysl. Zákazníci jsou rozmístěni po celém světě, ovšem největší zastoupení je v zemích Skandinávie, a to v soukromém i veřejném sektoru.

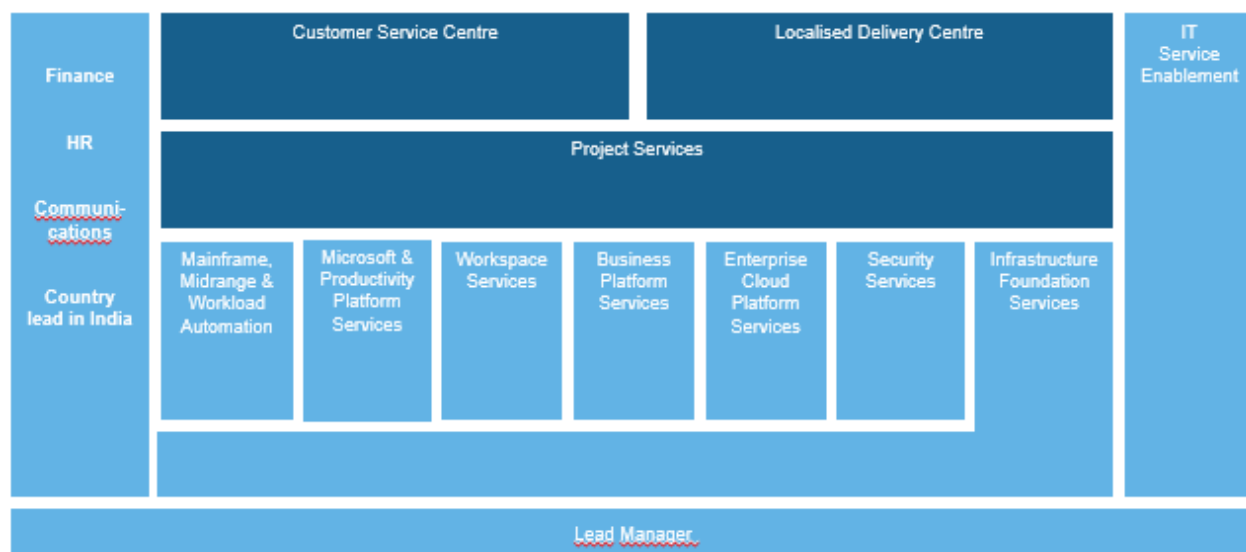
*Hodnoty definované pro práci v této společnosti jsou práce v týmech, péče o jejich zaměstnance i zákazníky, důraz na kvalitu, učení se a růst.*

### **3.1.2.Organizační struktura firmy**

Organizace má maticovou organizační strukturu vedení v jednotlivých zemích, průmyslových odvětvích a globálních službách. Jde o velmi komplexní a rozsáhlou organizační strukturu. Z širokého úhlu pohledu jsou na vertikální úrovni průmyslové skupiny, ty jsou tvořeny zákaznickými týmy a samotnými zákazníky. Horizontální úroveň je představována systémy pro integraci a konzultaci, řízenými službami, podpůrnými funkcemi a managementem. Samostatně vystupuje služba na vývoj produktů. Horizontální i vertikální úroveň se dále štěpí na subkategorie, subsubkategorie, atd., kde do každé kategorie vstupují jiná oddělení, služby, země, atd. Důležitými podpůrnými odděleními jsou např. marketing, finance, management, personalistika, ..., ale ty pro tuto práci nejsou stěžejní.

Pro diplomovou práci je hlavní pozornost věnována oddělení *Customer Service Centre* (dále CSC), které je součástí *MS Shared Services*, jak je možné vidět na Obr. 3.1. Cílem *MS Shared Services* je vyvíjet a dodávat mnohým zákazníkům standardní služby s vysokou kvalitou, automatizací, opakovatelností a cenovou konkurenceschopností. *MS Shared Services* je rozdělena do různých oblastí služeb, tzv. *Service Areas* (SA), které jsou po celou dobu odpovědné za poskytované služby. SA jsou důležité pro kompletní nabídku transformačního outsourcingu. Organizační struktura je rozdělena na SA, které jsou zodpovědné za poskytované služby (Microsoft, Cloud, atd.), a spravovány jsou službou základní infrastruktury. Nad těmito SA je služba projektů, neboť veškeré operace ve firmě jsou pomocí nich řešeny. Tyto oblasti jsou zastřešeny kontaktním místem pro zákazníka a místem pro speciální požadavky zákazníků. Jako podpůrné ale důležité služby jsou na jedné straně služby umožňující IT a na druhé straně služby finanční, personální a komunikační.

Celkově je pro firmu velmi důležité oddělení *MS Shared Services*, kde spadají všechna výše zmíněná oddělení, protože je pro firmu hlavním zdrojem finančních prostředků. Jde o řízené a sdílené služby, které jsou poskytovány zákazníkovi. Důraz je kladen na kvalitu, z čehož následně vyplývá spokojenost zákazníka.



Obr. 3.1 Organizační struktura řízených a sdílených služeb  
(Zdroj: interní stránky firmy + vlastní přepracování)

### 3.1.3. Procesy ve firmě

Všechny procesy ve firmě, jsou založeny na konceptu ITIL. Cartlidge a spol (2007) popisují ITIL jako veřejně dostupný rámec, jenž popisuje nejlepší praktiky ve Správě služeb IT. Poskytuje rámec pro zvládnutí IT v organizaci, pojednává komplexně o službách a zaměřuje se na neustálé měření a zlepšování kvality dodávaných služeb IT, a to jak z pohledu podnikání, tak z pohledu zákazníka.

Hlavními částmi pro správu služeb IT je *Strategie služeb*, *Návrh služeb*, *Přechod služeb*, *Provoz služeb* a *Neustálé zlepšování služeb*. Pro tuto práci je důležité brát v potaz *Provoz služeb*, neboť jde o doručování dohodnutých úrovní služeb uživatelům a zákazníkům, o správu aplikací, technologií a infrastruktury podporující dodávku těchto služeb. Je zde vytvářena skutečná hodnota podnikání.



### 3.2. Popis oddělení *Customer Service Centre*

Stavebními kameny CSC je *Global Service Operations Centre (GSOC)* a *Service Desk (SD)*, což jsou typické služby poskytované zákazníkům. Cartlidge a spol (2007) definuje *Service Desk* jako nezbytný nástroj pro správu informačních technologií (ITSM). Jde o komunikační centrum, jež poskytuje kontaktní místo (Single Point Of Contact – SPOC) mezi organizací a jejími zákazníky, zaměstnanci a obchodními partnery. Důležité je poskytnout zákazníkovi odpovídající pomoc v co nejkratším čase. Jeho hlavním úkolem je zaznamenávat všechny incidenty a požadavky včetně kategorizace a prioritizace, prozkoumávat a diagnostikovat je v první úrovni, spravovat celý životní cyklus incidentů a požadavků včetně jejich uzavírání zákazníkem schváleným řešením, průběžně informovat uživatele o stavu služeb, incidentů a požadavků.

CSC je rozdělen podle zemí na Švédsko a Finsko, Estonsko, Českou republiku. V těchto zemích jsou různě zařazeny podpůrné procesy - prodej, komunikace, vývoj. Pro zajištění stabilního chodu, kvality, optimalizace, standardizace služeb GSOC a SD je důležité zastoupení v podobě *Lead Product Manager*, *Product Manager / Standardization Specialist* a *Lead Service Manager*. Celé toto oddělení je pod vedením stanoveného manažera (Obr. 3.2). Z důvodů maticové organizační struktury je do CSC rovněž zahrnuto finanční a personální oddělení.



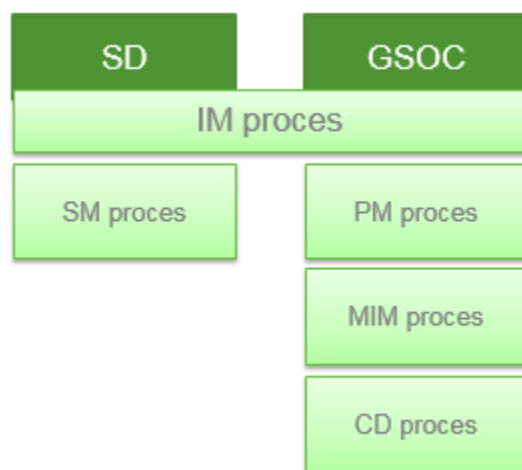
Obr. 3.2 Organizační struktura Customer Service Center  
(Zdroj: interní stránky firmy + vlastní přepracování)

Ve Finsku a Švédsku vede SD pouze jeden společný manažer (*Onshore SD*). V České republice (*Service Centre Czech*) a v Estonsku (*Service Centre Estonia*) vede tyto jednotky separátní manažer. Jednotlivé SD jsou dále rozděleny do dalších skupin, podle účelu SD – např. SD pro samotnou firmu jako zákazníka, SD pro externí zákazníky, ... Tyto oddělené jednotky jsou vedeny *Line* manažery, jež jsou zodpovědní za práci skupiny jako celku ale také za její specialisty.

SD je vedeno celkem 21 liniovými manažery, kdy jejich skupiny se početně liší (od sedmi do dvaceti specialistů). SD je celkově zastoupen 321 specialisty – 65 pro Finsko, 95 pro Švédsko, 97 pro Estonsko a 64 pro Českou republiku (čísla se měsíčně mění). Specialisti pracují na osmi hodinové směny s půlhodinovou pauzou (převážně od 7 do 15:30). Některé skupiny specialistů mají pouze denní směny, ostatním specialistům se směny střídají – denní, odpolední, noční. Podobné je to s prací přes týden a o víkendu.

### 3.2.1. Procesy v oddělení *Customer Service Centre*

Jak bylo v předešlé kapitole řečeno, jsou základními bloky CSC SD a GSOC. Nejen tyto dvě služby v rámci celého podniku jsou pokryty procesem *Incident Management* (IM). V rámci IM procesu je součástí každého bloku vymezen zvláštní proces – pro službu SD je to *Service Management* (SM) proces, pro službu GSOC je to proces *Problem Management* (PM), *MIM* a *Control Desk* (CD), jak je znázorněno na Obr. 3.3. Následující kapitoly práce jsou věnovány službě SD s procesy IM a SM.



Obr. 3.3 Procesy zastřešující služby SD a GSOC (Zdroj: vlastní zpracování)

Ke každému procesu firmy jsou nastavené určité metriky, které odrážejí realitu daného procesu. Oddělení si z každého procesu vybírají takové metriky, jež zachycují právě jejich poskytovanou kvalitu služby.

## 3.3. Proces Incident Management

Podle Cartlidge a spol (2007) je incident neplánované přerušení IT služby nebo snížení kvality služby IT. Pomocí procesu IM by měl být incident řešen okamžitě, a to pomocí trvalého,

dočasného nebo náhradního řešení. Je-li systémem detekován nebo zákazníkem SD specialistovi nahlášen incident (může jít také o problém<sup>6</sup> nebo změnu<sup>7</sup>, ale ty SD řešeny nejsou) je snahou v co nejkratším čase zajistit normální běh služby tak, aby byl pro operace podnikání minimalizován dopad.

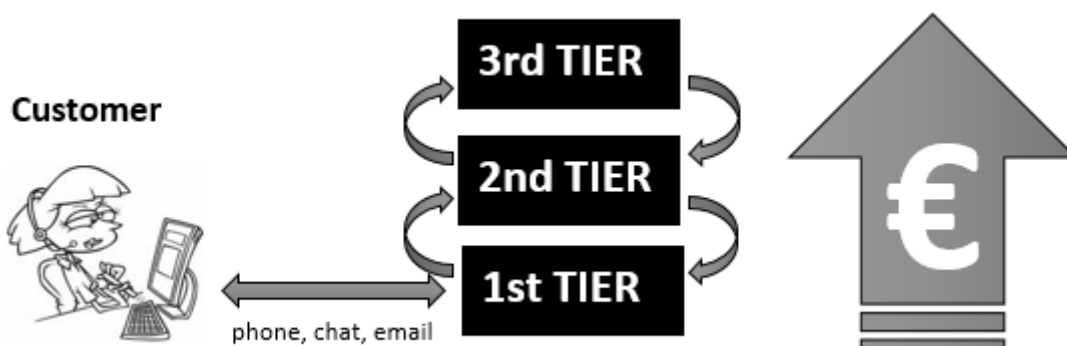
Zákazník pro jednání s SD využívá řadu komunikačních kanálů (*email, integrace, chat, telefon, atd.*), kdy je na základě incidentu nebo požadavku pomocí tiketovacího nástroje vytvořen tiket. Generování tiketů se liší od použitého komunikačního média. U každého tiketu je zaznamenán jeho kontaktní typ. Na základě emailu, chatu nebo telefonu je tiket SD specialistou zaveden do systému ručně. V případě integrace je tiket automaticky vygenerován systémem. Incident nemusí být nahlášen nebo vytvořen přímo zákazníkem, ale může vzniknout i pomocí nástrojů řízení událostí.

Není-li možné incident vyřešit rychle nebo SD specialisty, je nutné incidenty eskalovat - předat do vyšších úrovní. Jak je možné vidět na Obr. 3.4, tiket může být v rámci firmy řešen na třech úrovních. S danou úrovní je spjata i cena za tiket. Nejnižší úroveň (*1st tier*) je tvořena SD a cena tiketu je zde nejnižší. S kvalitou a rychlostí vyřešení tiketu (nejlépe na úrovni SD) úměrně roste i spokojenost klienta. S přeposíláním tiketů do vyšších úrovní (*2nd tier a 3rd tier*) roste, jak kvalifikace zaměstnanců, tak i cena za tiket. 2nd tier a 3rd tier je složen ze specialistů z různých oblastí IT (specialisté na operační systémy / servery, vývojáři / programátoři, specialisté na síť, BI specialisté, specialisté na reporting atd.). Podrobnější postup IM procesu je uveden v příloze 1.

---

<sup>6</sup> Problém může být důsledkem jednoho nebo více incidentů (incident je eskalován na problém), na které je při jeho vzniku odkazováno, nebo může vzniknout nezávisle na incidentu. Příčina je zjišťována v rámci *Problem management*.

<sup>7</sup> Jde o aplikaci jakékoli změny od původního stavu. Změna může být vyvolána zákazníkem, problémem, nebo automaticky vznikem událostí (např. změna o navýšení kapacity úložiště). Změny jsou spravovány procesem *Change management*.



Obr. 3.4 Proces Incident Management (Zdroj: vlastní zpracování)

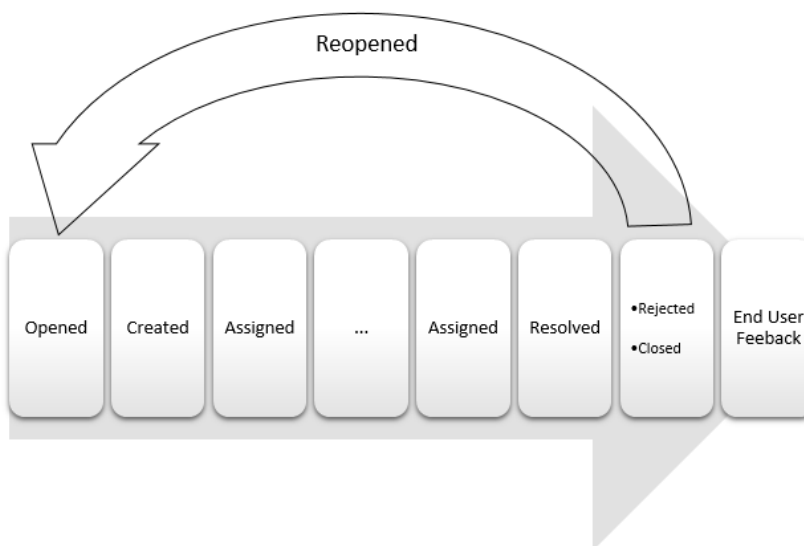
### 3.3.1. Životní cyklus řešení incidentu

Na Obr. 3.5 je znázorněn životní cyklus řešení incidentu. Nejdřív je incident otevřen (*Opened*) - jde o událost, kdy zákazník otevře formulář tiketu, vyplní a potvrdí ho. Je provedena identifikace, zda jde o incident. Poté následuje krok vytvoření (*Created*), kdy je serverem přijata informace o přichozím incidentu a je vytvořen záznam incidentu.

Incidentu je přiřazena služba, jež je k němu vymezena. Incident je zákazníkem / zaměstnancem ohodnocen podle dopadu, urgentnosti, na kterých je postavena priorita incidentu - *Low, Medium, High, Critical*. Od zvoleného ohodnocení a služby, jež zákazník v incidentu požaduje (vázáno na SLA), je stanovena maximální doba na vyřešení incidentu (může se měnit při přehodnocení incidentu).

Dále je incident předán (*Assigned*) specialistovi, který by ho měl vyřešit. Stav *Assigned* může incident nabývat několikrát během jeho životního cyklu – jde o předávání incidentu mezi specialisty firmy. Po nalezení řešení je toto řešení zapsáno do incidentu a je přepnut do stavu *Resolved*. Po této fázi rozhoduje o dalším stavu incidentu jen zákazník, ten ho může uzavřít (*Closed*), pokud je s řešením spokojen, nebo ho vrátí ke korekci zpět (*Rejected*). Incident je znovu otevřen (*Reopened*). Zákazník má sedm dnů, mezi fázemi *Resolved a Closed/Rejected*, aby se k řešení incidentu vyjádřil. Po sedmi dnech je incident automaticky uzavřen. Mezi stavy *Assigned a Resolved* může být incident v různých stavech, např. to může být *Awaiting Customer, Awaiting Specialist, Awaiting User Info, Work in Progress, Resolved Awaiting Verification* atd.

Po uzavření incidentu je zákazníkovi vygenerován dotazník *End User Feedback (EUF)*, ve kterém odpovídá na různé typy otázek – jak byl s řešením spokojen, jaká byla komunikace se specialisty, atd. Na základě těchto dotazníků může firma stále zvyšovat kvalitu poskytovaných služeb, aby byl zákazník spokojený a u této firmy zůstal.



Obr. 3.5 Životní cyklus incidentu (Zdroj: vlastní zpracování)

### 3.4. Současný stav zpracování a poskytování reportů

Cílem reportů tohoto oddělení je poskytovat strategické KPI o kvalitě poskytované služby, a to jak na úrovni jednotlivých specialistů SD, na SD skupinách, tak na globální úrovni služby jako takové. Do těchto strategických KPI, která jsou spjata s kvalitou služby, patří také EUF, který je klíčovým indikátorem, jelikož zaznamenává spokojenost se službou ze strany zákazníka. V součtu tato měření a výsledné reporty poskytují náhled na operativní část služby, a to na různých úrovních, tak, jak je zmíněno níže.

Snahou většiny manažerů je vytvářet takové metriky nebo předělávat výpočty tak, aby vždy vykazovaly dobrá čísla. Pro podnik tato varianta není vůbec přínosná, protože není vyvolán podnět na stálé vylepšování poskytovaných služeb, a organizace výkonností stagnuje, nebo v horším případě se zhoršuje. Metriky by měly být standardizované a automatizované, a měly by být vždy stanoveny hodnoty, kterých chce firma dosáhnout. V případě, že jich je dosaženo, mělo by dojít k navýšení / snížení dosahovaných hodnot, tak aby se výsledky nikdy dlouhodobě

nepohybovaly v požadovaných hodnotách. Pro firmu to v praxi znamená neustálou snahu o zlepšování a posouvání kvalit do vyšších úrovní.

### 3.4.1.Role v reportingu SD

Manažeři na různých úrovních podniku mají potřebu sledovat různé informace o dění v podniku. Při nerozdělení do skupin by pro některé manažery byly informace příliš strohé a pro některé zase příliš detailní. Někteří manažeři mohou být zastoupeni ve více rolích. Pro tuto práci byli manažeři rozděleni do tří rolí:

- **Manažeři vyšších úrovní**

Pro tuto úroveň manažerů je důležité sledovat výsledky za celé oddělení SD (převážně měsíční informace). Dělení je možné na úrovni zemí. Na této úrovni je možné počítat metriky komplexního charakteru jako *First Contact Resolution*, *First Level Resolution*, průměrný počet tiketů na 1 osobu, atd. Zde je zařazen tým vedení - manažer *MS Shared Service*, manažer *CSC*, Lead Service Manager, Finance and Sales Support.

- **Manažeři středních úrovní**

Pro manažery střední úrovně je typické monitorovat měsíční informace o jednotlivých skupinách SD v jeho zemi. Např. počet otevřených / vyřešených tiketů na skupinu, kolik tiketů bylo přeposlaných na skupinu. Do této skupiny manažerů jsou zařazeni manažeři jednotlivých zemí SD.

- **Manažeři nižších úrovní**

Manažeři nižších úrovní potřebují kontrolu nad službou na denní úrovni. Např. průměrný čas řešení tiketů, počet otevřených / vyřešených tiketů na člověka atd. Zde jsou zařazeni všichni *Line / Service managers* SD skupin, *Incident Managers*<sup>8</sup> a *Customer Incident Managers*<sup>9</sup>.

---

<sup>8</sup> *Incident Manager* – vede tým SD, je zodpovědný za bezchybný proces IM, eskaluje incidenty a požadavky, je zárukou přesného zpracování všech incidentů a požadavků na službu. Např. incident manažeři kontrolují specialisty IT oblastí, aby tikety byly vyřešeny, jak nejrychleji je to možné.

<sup>9</sup> *Customer Incident Manager* – jde o speciální tým manažerů sestavený přímo pro tuto firmu. Oproti *Incident Manager* je jeho pozice obrácena na funkčnost služeb poskytovaných zákazníkovi. Jeho úkolem je uspokojovat specifické potřeby zákazníků, řešit incidenty i urgentní, dohlíží na Service Level Agreement (SLA – dodatky ke smlouvám; ve firmě je pod touto zkratkou take chápán čas, do kdy musí být incident vyřešen), vytvářet zákaznické reporty, a další.

### 3.4.2.Poskytované metriky a reporty

V současné době je výše uvedeným manažerům vytvářena, sdílena a poskytována řada reportů a je vypočítáváno několik metrik. Ve většině případů velmi neefektivním způsobem. Mnoho reportů a metrik je komplexního charakteru a ze systému, s nímž uživatelé běžné pracují, nelze výsledky vyexportovat najednou. Proto zaměstnanci zodpovědní za poskytování reportů a metrik musí nejdříve vygenerovat separátní reporty, nad kterými jsou následně v MS Excel prováděny různé operace, ať už pomocí funkcí *MS Excel* nebo pomocí *Visual Basic for Application* (VBA). Mnohdy jsou data k měsíčnímu reportu stahována v několika separátních reportech (po několika dnech nebo týdnech), neboť z důvodů rozsáhlých dat není systémem umožněno vygenerovat report najednou - kvůli omezení nebo kvůli vytížení databáze.

Některé reporty jsou částečně standardizované – mají dané, jaké výsledky mají být reportem zobrazeny, jakou mají mít strukturu a vzhled. Ovšem tím, že jsou reporty z velké míry dodělávány ručně, neexistuje zde jistota stále správného výsledku (člověk vždy může něco přehlédnout nebo v některé fázi učinit chybný krok). U některých reportů není přesně stanovený postup k dosažení výsledku a někdy může být pozměněn, čímž vzniká problém s porovnáváním informací v trendu. Část reportů je vytvářena i na speciální požadavky, tzv. on-demand.

U KPI a metrik nastávají podobné komplikace, protože z velké části jsou KPI z výše zmíněných reportů dopočítávány. Správná definice a proces výpočtu je dán přesně, většinou podle již předem stanovených KPI a metrik organizací, které se tímto zabývají (např. MetricNet).

#### **Tiket report**

Primárním účelem tiket reportu je vyčíslit dotknuté tikety v jednotlivých zemích vzhledem k tomu, jestli byly uzavřeny v pracovní, nebo mimopracovní době. Princip spočívá ve vyexportování měsíčních uzavřených tiketů, jež byly otevřeny, vyřešeny nebo uzavřeny specialistou z SD. Jelikož tiket může být mezi jednotlivými SD přeposílán, je problém s jasným vytyčením země. Ta je vzorcem v MS Excel přiřazena v pořadí, zda byl daný tiket SD specialistou vyřešen, otevřen nebo zavřen. Dále je také pomocí vzorce zjištěno, jestli uzavření tiketu proběhlo v pracovní, nebo mimopracovní době. Mimo pracovní dobu jsou zahrnuty víkendy, státní svátky (tyto musí být zjištěny pro danou zemi, kde se SD nachází) a rozmezí od 16:00 do 8:00 v pracovních dnech. Pro jednotlivé země musí být brána v potaz i časová pásma. Tiket report je dále využíván i při počítání jiných metrik / KPI.

Nedostatkem tohoto reportu je, že není pevně nastavený proces, jak tento report vyexportovat. Zda mají být data pro report stahována vzhledem k SD specialistům nebo SD skupinám. Report má konečnou podobu až po spojení několika separátních reportů, protože jak už bylo výše řečeno, jde o velké množství dat a reportingové systémy mají omezené použití.

### **Report produktivity**

Cílem reportu produktivity je zachycovat výkonnost SD specialistů, SD skupin a SD podle zemí. Jde o komplexní report, kterým jsou sledovány otevřené / vyřešené tikety na výše zmíněné úrovni detailů. Počet vyřešených tiketů na SD skupiny je zachycován v měsíčním trendu a při přepočtení poměru zaměstnanců (zaměstnanci mohou řešit tikety v různých skupinách a pracovat na tiketech mohou i stážisté) je vypočten průměrný počet vyřešených tiketů na SD skupinu. Dále je v tomto reportu zahrnut počet unikátně dotčených tiketů, na kterých SD specialisté v daném měsíci pracovali, počet neunikátně dotčených tiketů (SD specialista může na stejném tiketu pracovat vícekrát) SD specialistou v měsíci. Monitorovány jsou i průměrné aktivní časy SD specialistů strávené při řešení tiketů a průměrné aktivní časy SD specialistů na „dotek“ tiketu.

Nedostatky tohoto reportu jsou více či méně podobné předchozímu. Zde však minimálně dva krát narůstá množství dat, která jsou potřebná pro zpracování tohoto reportu. Nárůst je způsoben tím, že je potřebné získat data pro SD specialisty a stavy tiketů z metrické tabulky, což je tabulka, v níž je zaznamenána veškerá aktivita na daném tiketu (průměrně může být množství záznamů až pětinasobně vyšší).

### **KPI / Metriky z Incident Management procesu**

#### *○ First Contact Resolution (FCR)*

FCR představuje procento manuálních incidentů vyřešených u kontaktního typu chat a telefon do hodiny a tiket neopustil původně kontaktované SD, u kontaktního typu email a integrace při prvním uložení. Je počítáno vzhledem ke všem tiketům otevřeným, vyřešeným a uzavřeným daným SD a stejných kontaktních typů. Roste-li hodnota tohoto KPI, roste i spokojenost zákazníka.

Tento ukazatel není nyní v podniku nijak automatizován. Za použití různých reportů a podnikových nástrojů je dopočítáván ručně. Sledován je měsíčně a za celé oddělení, v případě potřeby je možné sledovat za země nebo skupiny SD.



- *First Level Resolution (FLR)*

Je procento manuálních incidentů řešených pouze na úrovni SD (tikety mohly procházet mezi jednotlivými SD, ale do vyšších úrovní nebyly přeposlány) v porovnání se všemi tikety, které SD prošly. Hodnota tohoto ukazatele by měla být udržována ve vyšších hodnotách, neboť to znamená, že tikety jsou vyřešeny v nejnižších úrovních. To je pro firmu výhodné z hlediska nákladů.

Proces výpočtu FLR byl standardizován a zpracování bylo převedeno do *SAP Crystal Reports*, tudíž již nemusí být prováděn ručně, což šetří čas a výsledek je ihned k dispozici. Jde ale pouze o číslo zaznamenané v provozním revue, které je nanejvýš porovnáváno s minulou hodnotou. Zatím manažerům není automaticky přístupné, např. na dashboardu, ale reporting týmem je do revue v *MS PowerPoint* vypočtená hodnota zapsána a umístěna na sdílené úložiště.

- *Doteky tiketů*

Data pro tuto metriku lze brát z dotčených tiketů SD specialisty nebo SD skupinami. V podniku je v úvahu brána první varianta. Z pohledu celého oddělení je vypočten průměrný počet unikátních / neunikátních doteků na jednoho SD specialistu. K SD skupině, do které je SD specialista primárně přiřazen, může být dopočítána průměrná hodnota doteků na jednu skupinu.

Z hlediska vyšší detailnosti jsou sledovány unikátní / neunikátní doteky na konkrétní SD specialisty, tím pak mohou být sledovány počty i SD skupin. Hodnoty jsou v obou případech na měsíční bázi. Metrika je nyní částečně automatizovaná pomocí *SAP Crystal Reports*. Vygenerované číslo z hlediska oddělení jako celku je manažerům zapsáno do měsíčního provozního revue. Počty doteků ke konkrétním zemím, skupinám a specialistům SD jsou zobrazeny v pivot tabulce a data jsou manažerům posílána emailem.

- *Úroveň řešení*

Úroveň řešení SD je zachyceno procento vyřešených manuálních incidentů na SD v poměru se všemi incidenty, které v daném měsíci SD prošly. Touto metrikou je sledována úroveň řešitelnosti SD. Zkoumat ji je možné z různých úrovní detailu. Do provozního revue je hodnota řešitelnosti SD vkládána za všechny SD dohromady. Dále je tuto metriku možné agregovat na země nebo konkrétní SD skupiny.

- *Průměrný čas strávený na tiketu*

Jde o průměr durace na incidentech v aktivních stavech (*Assigned to, Work in Progress, Resolved-Awaiting Verification*) u daného SD specialisty nebo SD skupiny, vzhledem ke všem dotekům SD specialistou nebo SD skupinou. Jde o průměrný aktivní čas strávený na tiketu. Zkoumat lze z pohledu specialisty nebo skupiny. V nižších úrovních detailu metrika příliš nedává smysl.

Metrika je důležitá z pohledu utilizace specialisty (skupiny) a může být nápomocná k nalezení chyb uvnitř procesu, např. při příliš dlouhém průměrném čase. Všechny doteky SD specialistů a stavů, jsou na několikrát stahovány do MS Excel, kde pomocí VBA je ke každému SD specialistovi dopočten jeho průměrný aktivní čas. Nyní je metrika převedena do *SAP Crystal Reports*, kde již není zapotřebí do procesu výpočtu nijak zasahovat. Výsledky jsou manažerům ve formě tabulky posílány emailem.

- *End User Satisfaction*

End User Feedback představuje ohodnocení kvality služby zákazníkem, která mu byla poskytnuta při řešení incidentu nebo požadavku. Zákazníkovi je po několika vyřešených incidentech vygenerován dotazník, v němž vybírá z intervalu (1 až 4, kde 1 je špatné hodnocení) na čtyři položené otázky. Hodnota metriky představuje průměr vzhledem ke všem ohodnocením na úrovni celého oddělení SD, země SD nebo skupiny SD.

Metrika je ukazatelem kvality poskytované služby. Hodnota by se v nejlepším případě měla držet kolem čtyř, kterou je řečeno, že zákazník je se službou spokojen.

- *Opětovné otevření tiketů*

Úkolem opětovného otevření tiketů je ukázat procento tiketů, které musely být znovu otevřeny (řešení SD specialistů nebylo potvrzeno zákazníkem jako uspokojující) v poměru se všemi tikety, jež byly zavřeny. Procento znovu otevřených tiketů lze zkoumat za oddělení jako celek, za zemi, za skupinu a také za jednotlivé zaměstnance.

Metrika do teď nebyla nikde uveřejňována, ale nově bude použita v provozním revue. Jde o důležité číslo a jeho hodnota by se měla držet co nejbliž nule. Čím blíže bude nule, tím kvalitnější služba je doručována zákazníkovi.

Metrika zatím nebyla nijak standardizována ani automatizována. Výpočet čísla není příliš obtížný, ale je nutné opět provést několik kroků k jeho zjištění. V produkční databázi je uchovávána hodnota o počtech znovu otevření.

### **KPI / Metriky ze Service Management procesu**

#### *o Počet otevřených / vyřešených tiketů*

Počet otevřených / vyřešených tiketů je základní ukazatel SM procesu. Ukazatel je možné sledovat za celé oddělení SD, za země, za skupiny SD a za jednotlivé specialisty. Hlavně ukazatel vyřešených incidentů spolu s KPI FLR, jsou ukazateli produktivity. Může komplexně poskytovat informaci, jak úspěšně jsou incidenty řešeny a jak je služba kvalitní. Efektivní je sledovat tuto hodnotu na měsíční úrovni.

Získání dat pro report není příliš náročné a nezahrnuje ani složitý výpočet, hodnoty lze jednoduše vyčíslit pomocí pivot tabulky, v níž je možnost zobrazit výsledky na všech výše zmíněných detailech.

#### *o Překročené řešené / reakční SLA*

Překročené řešené SLA je měřené na SD skupinách a SD specialistech. Jde o počet tiketů, kde při jeho řešení nebylo dodrženo předem stanovené SLA a na řešení tiketu bylo potřeba více času. SD skupině nebo SD specialistovi je připsán tiket za podmínky, že v čase propadnutí měl tiket na sobě. Pro zajištění kvality dodávané služby by se hodnota měla pohybovat u nuly. Metriku je možné sledovat na denní, týdenní i měsíční bázi. Pro lepší přehlednost a rychlejší rozhodování, je vhodné mít tyto data denně pod kontrolou.

Překročené reakční SLA je SD skupině nebo SD specialistovi připočteno tehdy, pokud první reakce na tiketu nebyla zpracována ve specifikovaném reakčním čase SLA. U obou metrik nastává problém, že incident může být přeposlán těsně před propadnutím a incident propadne na někom, kdo incident již vyřešit nestihne. Výsledek pro obě metriky není obtížné zjistit pro SD skupiny, v databázi je uchovávána informace o skupině, na které tiket propadl. Pro SD specialisty tyto údaje vedeny nejsou a zjistit je lze pomocí specifického reportu.

#### *o Tikety poslané na vyšší Tier*

Jde o metriku, v které jsou zaznamenávány tikety přeposlané specifickým SD na 2nd nebo 3rd Tier (příp. i na jiné SD). Metrika je sledovaná na úrovni SD skupin a na měsíční bázi. Zjišťovat

metriku je možné, buď pro počty přeposlaných tiketů, nebo také míru přeposlání vzhledem ke všem tiketům, které daným SD prošly.

Pro vypočtení konkrétních čísel je potřeba sled několika komplexních kroků a běžný uživatel se bez specializované osoby k výsledku nedostane. Metrika zatím veřejně nebyla poskytována, ale je s ní počítáno do budoucích reportů (dashboardů). Metrika je důležitým ukazatelem procesu zpracování incidentu daným SD.

U všech výše zmíněných metrik a KPI je nevhodné jejich sdílení a přístup informací ze strany manažerů. Místo pro jejich zveřejňování není nijak automatizované ani standardizované. Některé metriky jsou pro vybrané manažery posílány emailem nebo ukládány na sdílené úložiště. Ostatní metriky jsou poskytovány na vyžádání.

### **3.4.3.Nástroje pro zpracování reportů**

Podkapitola je zaměřena na vytýčení a popsání nástrojů, pomocí kterých mohou být získávány nebo doručovány reporty pro oddělení.

#### **Informační systém ServiceNow**

Jde o komplexní systém firmy *ServiceNow*, který si firma podle svých potřeb upravila. Systémem je umožněna viditelnost do ITIL procesů. Je určen ke správě tiketů (tiketovací nástroj) a také zahrnuje řadu nástrojů ke správě jednotlivých ITIL procesů jako např. Incident Managementu, Problem Managementu, Knowledge Management, atd. Každému uživateli jsou zpřístupněny takové nástroje a možnosti v něm, jaká je mu v systému nastavena role. Data jsou do systému poskytována z produkční databáze.

Mimo jiné je součástí tohoto systému také nástroj pro reportování. Data pro reporty jsou čerpána, jak již bylo výše zmíněno z produkční databáze, v níž jsou např. tabulky zaznamenávající údaje o tiketech, o historii tiketů, o zákaznících, o zaměstnancích, o skupinách, o SLA, o konfiguračních položkách a o mnohém dalším. Výstupy reportů mohou mít několik podob, může jít o list záznamů, o kontingenční tabulku, grafy různých druhů (sloupcový, koláčový, ...), atd. Pro každý výstup jsou k dispozici specifické možnosti nastavení (pro list – výběr atributů, příp. grupování podle atributu, pro grafy – grupování a agregace podle zvoleného atributu). Primárně je možné pro zdroj reportů zvolit pouze jednu tabulku, z níž budou data čerpána. V attributech je poté

na výběr i z atributů jiných tabulek, které jsou s ní ve spojení. Nad daty jsou zaváděny různé filtry, jež jsou určeny pro vymezení potřebné skupiny dat.

Jde o statické reporty, které mohou být případně časově naplánovány a zaslány emailem. V tomto nástroji neexistuje žádná možnost provedení výpočtu, úpravy či připojení další tabulky. Další nevýhodou je omezené exportování záznamů – padesát tisíc řádků nebo pět set tisíc buněk a sto dvacet sekund pro natažení dat z databáze, což je při velkých reportech nedostačující. Reporty mohou nastavovat specialisti na reporting, ale mohou si je nastavovat a exportovat uživatelé, jež mají umožněno s tímto nástrojem pracovat.

### **SAP Crystal Reports**

*SAP Crystal Reports* od firmy *SAP* je reportingový nástroj fungující na principech BI. Nástrojem je umožněna kombinace různých zdrojů. Jde o jednodušší a intuitivnější přístup k datům. SW povoluje používat funkce jako *Drag-and-drop*, *Drill-down*, atd. Data lze zobrazit v listu, v grafu, v mapě, lze nastavovat grupování, agregace a různě přizpůsobovat reporty. Jde o velmi užitečný SW, který by spolu se *SAP Business Objects* (popsáno níže) mohl pro uživatele sloužit jako kvalitní nástroj pro zveřejňování reportů (dashboardů), kde by si uživatelé předpřipravené reporty mohli doladit dle vlastních potřeb.

Se SW pracují specialisti na reporting. Tabulky produkční databáze je možné různě kombinovat pomocí odpovídajících atributů. Zákazníkům a zaměstnancům jsou vytvořeny statické reporty, maximálně s funkcí *drill-down*. Reporty jsou zmíněným osobám, jednotkám poskytovány většinou ve formě tabulkového souboru, a jsou buď posílány emailem, nebo nahrávány na sdílené úložiště, kde si je uživatelé mohou stáhnout. V reportech není žádná interaktivita a flexibilita.

### **Microsoft Excel**

Jde o tabulkový procesor od firmy Microsoft. Ve firmě je nejpoužívanější SW pro sledování a zaznamenávání potřebných údajů. Mnoho uživatelů (hlavně interní zaměstnanci) si stahují data z výše zmíněných nástrojů a vkládají do Excelu (buď je to nutné, nebo potřebují provádět úpravy nad staženými daty). Je využíváno funkcí Excelu či VBA. Přepracované reporty jsou poté prezentovány, sdíleny na veřejná úložiště nebo jsou potřebná čísla přepisována do prezentačního nástroje, kdy až tento je např. sdílen.

### 3.4.4. Procesy vytváření metrik a lidské zdroje

Vytvářené reporty a procesy jejich přípravy nebyly ve firmě nikdy standardizovány. Reporty jsou a byly nastavovány na míru každému, kdo reporty potřeboval / potřebuje. Vzhled reportu je více či méně daný (hlavička, barvy, atd. – může se lišit vzhledem k zákazníkům, nebo k zaměstnancům firmy), ovšem obsah je přizpůsobován na míru uživatele, nemluvě o poskytování reportů (popsáno níže).

Ve firmě je vytýčeno oddělení pro reporting, ale tím že zaměstnanci mají přístup do reportingového nástroje v *ServiceNow*, mohou reporty vytvářet nezávisle na reportingovém týmu. Tím mohou vznikat nekompatibilní a protichůdné výsledky poskytovaných reportů manažerům, či jiným uživatelům. Reportům vlastní potřeby nesdílené jiným osobám nemusí být věnována taková pozornost, ale i tak uživatelé mohou generovat reporty, které byly již dříve vytvořeny – duplikace reportu a dvojnásobné zatížení databáze. Reporty prezentovány dalšímu okruhu lidí by měly být před vytvořením konzultovány (od koho a jaké reporty již obdrží; co má být jejich výsledkem), aby nedocházelo již k výše zmíněným nedopatřením.

### 3.4.5. Zdroje pro reporty

Cílem této podkapitoly je specifikovat zdroje, která mohou sloužit jako základna pro reporty, metriky a KPI.

#### **Produkční databáze**

Jde o relační databázi na platformě *MySQL*, jež má pro *ServiceNow* dané propojení jednotlivých tabulek. V *SAP Crystal Reports* je možné si vybrat jen potřebné tabulky a rovněž si je podle potřeby propojit.

Databáze je složena z několika set tabulek. Pro tuto práci jsou důležité tabulky, které zaznamenávají informace o tiketech a o jejich historii, o zaměstnancích, o skupinách, o zákaznících, o ohodnoceních.

#### **Enterprise Resource Planning**

Snahou *Enterprise Resource Planning* (dále ERP) systému od firmy *ORACLE* je optimalizovat finanční procesy s cílem usnadnit lepší rozhodování, transformovat řízení výdajů vzhledem k maximalizaci hodnoty zdrojů a snížení manuálních procesů.

V databázi jsou uchovávány podrobné údaje o zaměstnancích – jméno, příjmení, email, pozice, manažer, zařazení v organizační struktuře. Oproti produkční databázi, v níž jsou uloženy pouze aktuální údaje o daném zaměstnanci, je v této databázi obsažena i historie změn pro každého zaměstnance (např. změna manažera, organizační jednotky, pozice). Jsou v ní zachycovány údaje o odpracovaných hodinách, o účinnosti na projektech, o dovolených, o navštívených projektech, atd.

#### **Databáze hovorů**

Jde o zvláštní databázi zaznamenávající čas hovorů na jednotlivých jednotkách, přeměrování hovorů, úspěšné nebo neúspěšné ukončení hovorů.

### **3.4.6. Možnosti dodávání reportů / metrik**

Reporty mohou být uživatelům dodávány různými způsoby, jakými je shrnuto v této podkapitole.

#### **SharePoint**

SharePoint v této organizaci slouží jako nástroj pro vytváření a sdílení dokumentů a informací pro definované pracovní skupiny firmy. Stránka je dostupná pouze pro interní uživatele a do každé skupiny jsou přidělováni jednotliví uživatelé s právy. Umožňuje přímou správu dokumentů.

Uživatelům jsou do vybraných skupin a složek ukládány potřebné reporty, které byly po předchozí domluvě nastaveny.

#### **SAP Business Objects**

*SAP Business Objects* je BI systém na podporu podávání reportů, informací, KPI uživatelům – zaměstnancům, zákazníkům či partnerům. Pomocí různých nástrojů je možné výsledky vizualizovat do takové podoby, která je snadno pochopitelná a přehledná. V jeho nabídce je také poskytování ad-hoc reportů.

V organizaci je tento systém používán jako webové úložiště a automatický spouštěč reportů posílaných emailem např. na denní, týdenní, měsíční bázi. Reporty jsou vizuálně a obsahem dat předpřipraveny v *SAP Crystal Reports*. Tyto reporty jsou rozděleny do tří částí, podle toho kde a jak jsou zveřejňovány uživatelům. Jedna část má uloženou instanci v *SAP Business Objects*, kterým je časově naplánováno odesílání reportů (ve formátu *pdf*, *xls*) na emailové adresy uživatelů.

Druhá část má zde rovněž uloženou instanci, v níž je nastaveno automatické spouštění reportů dle nastavených podmínek a parametrů, rozdíl však spočívá v ukládání vygenerovaných instancí (ve formátu *rpt* – *Crystal Reports*) do tohoto webového úložiště s funkcionalitou *drill-down*. I když jsou v těchto reportech obsaženy parametry pro vyfiltrování specifické skupiny dat, jsou ve většině případů reporty nastaveny např. pro konkrétního zákazníka se standardní množinou hodnot parametrů. Třetí část je ukládána do *ReportView*.

### **ReportView**

Opět jde o webové úložiště, kde je sdílána třetí část reportů vytvořených v *SAP Crystal Reports*. Pomocí různých aplikací vyvinutých zaměstnanci podniku jsou automaticky vygenerované reporty uloženy do tohoto prostředí (ve formátu *pdf* nebo *xls*).

### **3.4.7. Shrnutí**

Největším problémem pro vytváření strategických KPI není množství zdrojů, z kterých jsou data čerpána, ale různé nástroje pro zpracování reportů, nestandardizované procesy pro výpočty a také nejednotné úložiště pro poskytování KPI.

Nevýhodou organizace je používání značné škály nástrojů, u kterých ve většině není využito všech funkcí, jež je daným nástrojem nabízeno. Rovněž jde i o finanční stránku, neboť firma platí za nástroje / systémy pokrývající jen určitou oblast, namísto, aby důkladně aplikovala nástroje / systémy, v nichž budou zahrnuty všechny možné oblasti. Všechny nástroje, jež byly v této práci zmíněny, byly popsány ze strany interního nástroje pro práci a reporting. Nástroje nebyly popisovány jako implementovaná služba zákazníkům.



## 4. Návrh řešení a implementace

V této kapitole jde o praktickou realizaci diplomové práce. Je zde navrhována reportingová platforma pro IT oddělení. Na základě životního cyklu dimenzionálního modelování od R. Kimballa je navržen dimenzionální model, dále je proveden fyzický návrh dimenzionálního modelu, jsou navrženy datové pumpy pro nahrávání dat do jednotlivých dimenzí a tabulky faktů. V poslední fázi následuje návrh dashboardů na principech User Experience.

### 4.1. Dimenzionální modelování

Podle výše zmíněného životního cyklu je dimenzionální modelování uskutečněno ve čtyřech krocích, a to zmapování podnikových procesů a požadavků, stanovení granularit faktové tabulky, identifikace dimenzí a faktů. Po tomto návrhu bude sestrojen konceptuální a logický model.

#### 4.1.1. Zmapování podnikových procesů a požadavků

Vzhledem k teoretické části by v této fázi měly být vyspecifikovány veškeré procesy firmy, ohodnoceny dle různých druhů faktorů (složitost systémů, dostupnost a kvalita dat, strategický význam pro podnik), měla by být stanovena priorita ke stanovení pořadí zpracování. V této fázi to nebude zapotřebí, neboť byl vybrán konkrétní proces pro IT oddělení a je zpracováván tento konkrétní proces.

V třetí kapitole *Analýza současného stavu zpracování a poskytování reportů v IT oddělení* bylo řečeno o procesech firmy jako celku (vycházející z koncepce ITIL – *Incident Management*, *Change Management*, *Problem Management* a mnoho dalších procesů nastavených firmou nebo, dáno ITILEm). Při zadávání diplomové práce bylo vymezeno oddělení i procesy IT firmy, pro které bude vytvářen dimenzionální model. Hlavní pozornost je věnována IM procesu a jeho podprocesu SM. Složitost zdrojových systémů IM byla označena jako střední. Dostupnost dat ze zdrojových systémů je také na střední úrovni – data v databázích jsou dostačující, není zapotřebí externích zdrojů, ovšem komplikace nastávají s bezpečností a přístupovými právy. Kvalita dat ve zdrojových systémech je velmi dobrá. Sledování IM procesu na úrovni SD má pro podnik vysoký význam.

Proces IM byl popsán v kapitole 3.3 a jeho podrobné schéma postupných kroků (neboli workflow) je uveden v příloze 1. SM proces je zde zahrnut jako podrobné zkoumání práce

jednotlivých zaměstnanců SD, SD skupin a SD jako celku. Zahrnuje komplexní pohled na řešení incidentů.

V tabulkách Tab 4.1 a Tab 4.2 je sestrojen sémantický model pro oba procesy.

<b>Zkoumaný proces:</b>	P1: Incident Management Proces
<b>Úroveň podrobnosti:</b>	incidenty na denní bázi
<b>Návrh dimenzí:</b>	čas
	incident
	zákazník
	zaměstnanec
	skupina
	služba
	konfigurační položka
<b>Návrh metrik:</b>	počet otevřených / vyřešených incidentů
	znovuotevření incidentů
	hodnocení koncových uživatelů/zákazníků.
	udržení SLA

*Tab. 4.1 Sémantický model IM procesu (Zdroj: vlastní zpracování)*

<b>Zkoumaný proces:</b>	P2: Service Management Proces
<b>Úroveň podrobnosti:</b>	průchody incidentů na denní bázi
<b>Návrh dimenzí:</b>	čas
	incident
	zaměstnanec
	skupina
<b>Návrh metrik:</b>	First Contact Resolution
	First Level Resolution
	doteky tiketů
	úroveň řešení
	průměrný čas strávený na tiketu
	přeposlání na vyšší tier

*Tab. 4.2 Sémantický model SM procesu (Zdroj: vlastní zpracování)*

### Bus matrix

V následující tabulce Tab. 4.3 jsou pomocí bus matrix znázorněny procesy a k nim odpovídající měřitelné ukazatele a dimenze pohledů. Bude navržen jeden dimenzionální model s maximálně dvěma tabulkami faktů.

Proces	Granularita záznamů	Měřitelné ukazatele	čas	zákazník	zaměstnanec	skupina	služba	konfigurační položka	incident
P1: IM	počet incidentů za den	viz Tab. 4.1	x	x	x	x	x	x	x
P2: SM	počet za den přeposlaných incidentů	viz Tab. 4.2	x		x	x			

*Tab. 4.3 Bus matrix (Zdroj: vlastní zpracování)*

Na základě velké podobnosti a podmnožiny SM procesu bylo rozhodnuto, že dojde ke spojení obou procesů a v dalších částech práce bude navržena jedna tabulka faktů s podrobností záznamu na úrovni průchodů tiketů.

### Identifikace zdrojů

V této fázi dimenzionálního modelování měly být identifikovány zdrojové systémy dat důležité k měření výkonnosti a kvality vymezeného procesu. Tyto zdroje již byly popsány v kapitole 3.4.5 *Zdroje pro reporty*. Primárním zdrojem pro čerpání dat do dimenzionálního modelu je produkční relační databáze běžící na platformě MySQL. Jde o centrální úložiště firmy o několika set tabulkách a vzájemných propojeních. V práci bude použita jen část těchto tabulek.

Na žádost firmy je využita zkušební část dat převedena do SQL Server 2012, neboť v produkční databázi jsou uloženy citlivé údaje o zaměstnancích, zákaznících, službách, atd., jež nemohou být v externím dokumentu prezentovány. Převedení zkušebních dat je i kvůli zatížení systému, jež by při řešení mohl nastat. Struktura těchto dvou databází je stejná, rozdíl je ve zkušebních datech.

### Sběr požadavků

Sběr požadavků je velmi zásadní činnost, a to proto, aby byly nalezeny veškeré požadavky, které v budoucích reportech / dashboardech bude zapotřebí sledovat. Rovněž je důležité, aby uživatelská přání byla pochopena a nedošlo k záměně myšlenek.

Vzhledem k pracovní pozici byla ke sběru požadavků využita konzultace s vybranými manažery a konzultantem na KPI, kde bylo diskutováno, co se nyní měří, co by se mělo měřit, co

se skutečně bude měřit a jak se to má měřit. K tomu bylo využito již stávajících zdrojů, aby nebylo opomenuto na důležité metriky a provedla se předběžná konsolidace. Nejrozsáhlejší spolupráce probíhala právě s konzultantem na KPI, který je k tomu určený, a byl vytvořen následující seznam požadavků (Tab. 4.4). Všechny požadavky byly ohodnoceny prioritou 1 – 5, kde 5 má vysokou prioritu.

Č	Požadavek	Vybraná metrika	Entity	Priorita
1.	Kolik incidentů bylo otevřeno/vyřešeno/zavřeno SD skupinami / SD specialisty pro jakého zákazníka, službu nebo konfigurační položku s jakou prioritou, kontaktním typem? (manažer)	počet otevřených/ vyřešených/ zavřených incidentů	čas, skupina, zaměstnanec, zákazník, služba, konfigurační položka, incident	5
2.	Poměr vyřešených tiketů k otevřeným na SD skupinách / SD specialistech.	poměr vyřešených incidentů k otevřeným	čas, skupina, zaměstnanec	4
3.	Kolik tiketů propadlo na SD skupině / SD specialistovi v porovnání také s prioritou a zákazníkem? (manažer)	počet propadlých incidentů	čas, skupina, zaměstnanec, zákazník, incident	5
4.	Jak SD skupiny a SD jako celek udrželi stanovené SLA v poměru se všemi vyřešenými tikety? (manažer)	udržení SLA	čas, skupina	5
5.	Jak jsou zákazníci spokojeni s SD a SD skupinami při vyřešení incidentů? (manažer)	spokojenost zákazníků	čas, skupina	4
6.	Kolik z vyřešených incidentů SDs jako celku / SD skupinami / SD specialisty muselo být znovu otevřeno / + poměr s vyřešenými incidenty? (manažer)	počet znovuotevřených incidentů	čas, skupina, zaměstnanec, zákazník, incident	4
7.	Jaká je úroveň znovuotevření incidentů SD jako celku / SD skupin / SD specialistů (poměr znovuotevřených incidentů k vyřešeným)? (manažer)	úroveň znovuotevření	čas, skupina, zaměstnanec, zákazník, incident	4
8.	Kolika doteků bylo na daném SD specialistovi a SD skupině /+ průměr za SD jako celek? (konzultant)	počet doteků	čas, skupina, zaměstnanec	4
9.	Jaká je úroveň řešení jednotlivých SD skupin / SD specialistů / SD jako celku (poměr doteků k vyřešeným incidentům)? (konzultant)	úroveň řešení	čas, skupina, zaměstnanec	4
10.	Kolik incidentů SD skupina přeposlala na vyšší tier / + poměr se všemi doteky SD skupiny? (konzultant)	počet přeposlaných incidentů / úroveň přeposlání	čas, skupina, incident	3
11.	Jaký je průměrný čas SD specialisty / SD skupiny strávený na tiketu? (konzultant)	průměrný čas strávený na tiketu	čas, skupina, zaměstnanec, incident	3

12.	Kolik tiketů bylo vyřešeno jen na první úrovni / + poměr se všemi tikety, kterých se nějaké SD dotklo (ne průchody) nebo jen s počty tiketů, které na SD začaly - tudíž je mohli vyřešit. (konzultant)	počet tiketů vyřešených na první úrovni / FLR	čas	3
13.	Kolik tiketů (chat, phone, email, integrace) bylo vyřešeno při první kontaktování SD / + poměr s vyřešenými tikety (chat, phone, email, integrace) SD. (konzultant)	počet tiketů vyřešených při prvním kontaktování / FCR	čas, skupina, zaměstnanec	3

Tab. 4.4 Tabulka požadavků s danou prioritou (Zdroj: vlastní zpracování)

Pro oddělení je důležité sledovat všechny metriky, proto mají všechny výše zmíněné požadavky nastavenou relativně vysokou prioritu. Jednotlivá pozorování budou sledována na následující úrovni sumarizace vzhledem k časové jednotce:

- denní pozorování – požadavek č. 4, 7,
- týdenní pozorování – požadavek č. 1, 2, 3, 4, 5, 7, 8,
- měsíční pozorování – všechny požadavky.

Jak je možné vidět, jsou vyžadovány tři různé sumarizace, proto bylo rozhodnuto, že v tabulce faktů budou data uchovávána na nejnižší úrovni detailu. Agregace bude uskutečněna až ve výsledných dashboardech.

#### 4.1.2.Stanovení granularity

Tato kapitola je zaměřena na stanovení granularity – úrovně podrobnosti v tabulce faktů – co je každým řádkem v tabulce faktů zachyceno.

Diplomová práce je zaměřena na proces IM a jeho podproces SM, kdy je pro oba dva procesy dostačující využít jednu tabulku faktů s granularitou uvedenou v tabulce Tab. 4.5.

Fakta	1 řádek	Granularita
Day Revue	1 řádek = průchod tiketu	Průchod tiketů skupinou s odpovídajícími specialisty a stavy tiketů, spolu s nadcházejícími stavy a skupinami, vzhledem s podrobností k zákazníkovi, službě, konfigurační položce a incidentu.

Tab. 4.5 Granularita tabulky faktů (Zdroj: vlastní zpracování)

Do tabulky faktů *Day Revue* jsou data ukládána s vysokou granularitou, která odpovídá úrovni podrobnosti v produkční databázi, jde o tzv. *transakční typ tabulky faktů* (kapitola 2.3.6 *Identifikace faktů*).

Ke stanovené granularitě daného procesu je možné předběžně určit dimenze a fakta. Jde o dimenze – skupina, uživatel, zákazník, služba, konfigurační položka, charakteristiky tiketu, a potencionální fakta – otevřené, vyřešené, zavřené, propadlé, znovuotevřené tikety, spokojenost uživatelů, průchody tiketů, tikety vyřešené na SD, na první úrovni, při prvním kontaktování, tikety přeposlané na vyšší tier.

### 4.1.3. Identifikace dimenzí

Z již dříve provedených analýz, sběru požadavků, stanovení granularity a jejich předběžných dimenzí a faktů, je nyní možné lépe a přesně specifikovat všechny potřebné dimenze. Vzhledem k jednoduchému dimenzionálnímu modelu není zapotřebí využít žádné sdílené dimenze (ta je určena k propojení celopodnikového datového skladu, či jednotlivých tabulek faktů).

#### Návrhy dimenzí

Pro tento dimenzionální model jsou navrženy tyto dimenze:

- User,
- Group,
- Customer,
- Service,
- Configuration Item,
- Ticket.

V tabulkách Tab. 4.6 – 4.11 jsou vypsány jednotlivé atributy dimenzí, zdroj čerpání a pravidla pro datovou pumpu. Do každé dimenzi je zařazen umělý klíč v podobě automatického přírůstku o jedničku (projeví se u přidávání nových záznamů), jež je určen k jednoznačnému identifikování záznamu v tabulce faktů. V každé dimenzi je rovněž obsažen produkční klíč ze zdrojové tabulky.

<b>Dimenze:</b>	User		
<b>ID:</b>	Dim_User		
<b>Cíl</b>		<b>Zdroj</b>	
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>
userID	userID	-	PK, auto-increment
IDuser	idUser	<b>sys_user{sys_id}</b>	ZP
Login	login	<b>sys_user{u_customer_id}</b>	
User name	user_name	<b>sys_user{name}</b>	
User country	u_country	<b>sys_user{country}</b>	
Email	email	<b>sys_user{user_name}</b>	
User unit	user_unit	sys_user{company} -> core_comapny{sys_id} => <b>core_company{name}</b>	
User top company	user_top_company	sys_user{company} -> core_comapny{sys_id}, core_company{u_top_company} -> core_company{sys_id} => <b>core_company{name}</b>	
Primary group	primary_group	sys_user{u_primary_group} -> sys_user_group{sys_id} => <b>sys_user_group{name}</b>	
Manager email	manager_email	sys_user{manager} -> sys_user {sys_id} => <b>sys_user{user_name}</b>	
Manager name	manager_name	sys_user{manager} -> sys_user {sys_id} => <b>sys_user{name}</b>	
Active	active	<b>sys_user{active}</b>	

*Tab. 4.6 Dimenze uživatele (Zdroj: vlastní zpracování)*

<b>Dimenze:</b>	Group		
<b>ID:</b>	Dim_Group		
<b>Cíl</b>		<b>Zdroj</b>	
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>
groupID	groupID	-	PK, auto-increment
IDgroup	idGroup	<b>sys_user_group{sys_id}</b>	ZP
Group name	group_name	<b>sys_user_group{name}</b>	
Group type	group_type	<b>sys_user_group{u_group_type}</b>	
Group country	group_country	<b>sys_user_group{u_country_value}</b>	
Group unit	group_unit	sys_user_group{u_incident_manager} -> sys_user{sys_id}, sys_user {company} -> core_company{sys_id}= > <b>core_company{name}</b>	
Group top company	group_top_company	sys_user_group{u_incident_manager} -> sys_user{sys_id}, sys_user {company} -> core_company{sys_id}, core_company{u_top_company} -> core_company{sys_id} => <b>core_company{name}</b>	
Active	active	<b>sys_user_group{active}</b>	

*Tab. 4.7 Dimenze skupiny (Zdroj: vlastní zpracování)*

<b>Dimenze:</b>	Customer		
<b>ID:</b>	Dim_Customer		
<b>Cíl</b>		<b>Zdroj</b>	
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>
customerID	customerID	-	PK, auto-increment
IDCustomer	idCustomer	<b>core_company{sys_id}</b>	ZP
Customer name	customer_name	<b>core_company{name}</b>	
Customer top company	customer_top_company	core_company{u_top_company} -> core_company{sys_id} => <b>core_company{name}</b>	

*Tab. 4.8 Dimenze zákazníka (Zdroj: vlastní zpracování)*



<b>Dimenze:</b>	Service			
<b>ID:</b>	Dim_Service			
<b>Cíl</b>		<b>Zdroj</b>		
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>	
serviceID	serviceID	-	PK, auto-increment	
IDService	idService	cmdb_ci{sys_id}	ZP	cmdb_ci{sys_class_name}='service_offering'
Service name	service_name	cmdb_ci{name}		
Service company	service_company	cmdbi_ci{company} -> core_company{sys_id} => <b>core_company{name}</b>		
Operational status	operational_status	cmdb_ci{operational_status}		
Lifecycle	lifecycle	cmdb_ci{install_status}		

Tab. 4.9 Dimenze služby (Zdroj: vlastní zpracování)

<b>Dimenze:</b>	Configuration Item			
<b>ID:</b>	Dim_ConfigurationItem			
<b>Cíl</b>		<b>Zdroj</b>		
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>	
configurationItemID	configurationItemID	-	PK, auto-increment	
IDci	idCI	cmdb_ci{sys_id}	ZP	
CI name	ci_name	cmdb_ci{name}		
CI class	ci_class	cmdb_ci{sys_class_name}		
CI subcategory	ci_subcategory	cmdb_ci{subcategory}		
CI company	ci_company	cmdbi_ci{company} -> core_company{sys_id} => core_company{name}		
Operational status	operational_status	cmdb_ci{operational_status}		
Lifecycle	lifecycle	cmdb_ci{install_status}		

Tab. 4.10 Dimenze konfigurační položky (Zdroj: vlastní zpracování)

<b>Dimenze:</b>	Ticket		
<b>ID:</b>	Dim_Ticket		
<b>Cíl</b>		<b>Zdroj</b>	
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>
ticketID	ticketID	-	PK, auto-increment
IDticket	idTicket	task{sys_id}	ZP
Number	number	task{numer}	
Task type	task_type	task{contact_type}	
Category	category	task{category}	
Contact type	contact_type	task{contact_type}	
Priority	priority	task{priority}	
State	state	task{state}	
SLA solving	sla_solving	task{u_sla_solving}	
SLA reaction	sla_reaction	task{u_sla_reaction}	
Opened	opened	task{opened_at}	"dd/MM/yyyy"
Resolved	resolved	task{resolved_at}	"dd/MM/yyyy"
Closed	closed	task{closed_at}	"dd/MM/yyyy"
Active	active	task{active}	

*Tab. 4.11 Dimenze tiketu (Zdroj: vlastní zpracování)*

### Granularita v tabulkách dimenzí

V tabulce Tab.4.12 je popsána granularita jednotlivých dimenzí.

<b>Dimenze</b>	<b>Granularita dimenze</b>
User	Informace o uživateli včetně primární skupiny, jednotky a manažera.
Group	Informace o skupinách včetně země a jednotky.
Customer	Informace o zákaznících včetně nadřazené domény.
Service	Informace o poskytovaných službách včetně přiřazeného zákazníka.
Configuration Item	Informace o všech konfiguračních položkách (HW, SW) včetně sub-kategorie.
Ticket	Informace o všech tikech včetně typu a kategorie tiketu, kontaktního typu.

*Tab. 4.12 Granularita v tabulkách dimenzí (Zdroj: vlastní zpracování)*

### Dimenze data

Jde o speciální dimenzi, jak již bylo řečeno v teoretické části. Pro tuto práci je dostačující dimenze data, jeho atributy je možné vidět v tabulce Tab. 4.13. Data pro tuto dimenzi jsou čerpána z externího souboru (.csv), kde je zajištěno dostatečné množství dat pro nadcházející využívání dimenzionálního modelu.

<b>Dimenze:</b>	Calendar		
<b>ID:</b>	Dim_Calendar		
<b>Cíl</b>		<b>Zdroj</b>	
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>
dateID	dateID	-	PK, auto-increment
Date	date	externí soubor (.csv)	ZP; "dd/MM/yyyy"
Weekday	weekday		
Day name	weekday_name		
Day in month	day_in_month		
Day in year	day_in_year		
Week	week		
Month	month		
Month name	month_name		
Qtr	qtr		
Year	year		
Half	half		
Year_qtr	year_qtr		

*Tab. 4.13 Dimenze data (Zdroj: vlastní zpracování)*

### Hierarchie v tabulkách dimenzí

V této fázi je identifikována hierarchie (vztah předek – potomek) mezi jednotlivými atributy v tabulkách dimenzí. Nalezené vazby jsou sepsány v tabulce Tab. 4.14.

Dimenze	Popis hierarchie
User	User country → User Top company → User unit → Primary group → Manager name → User name
Group	Group country → Group Top company → Group unit → Group type → Group name
Customer	Customer Top company → Customer name
Service	Service company → Service name
Configuration Item	CI company → CI class → CI name → CI subcategory
Ticket	Task type → Category
Calendar	1. Year → Half → Qtr → Month name (Month) → Week → Weekday name (Weekday) → Date 2. Year-qtr → Month → Week → Day in year → Date 3. Year → Month → Day in month → Date

*Tab. 4.14 Hierarchie v tabulkách dimenzí (Zdroj: vlastní zpracování)*

## Změny v tabulkách dimenzí

Jde o vytýčení pomalu a rychle měnících se dimenzí a určení jejich typů, podle toho, jak často dochází ke změně atributů v dané dimenzi. O vyspecifikovaných dimenzích v předešlé části je možné říci, že se jedná o dimenze pomalu měnící (SCD), a to typu číslo jedna, u kterých dochází pouze k přepsání daného záznamu. V tomto dimenzionálním modelu se nenachází žádná rychle se měnící dimenze (RCD). Možné změny u SCD typu 1:

- *User* – u dimenze uživatele může dojít ke změně příjmení – změna ale není příliš častá; při změně pozice zaměstnanců může být pozměněna primární skupina, jednotka, jméno a email manažera; při odchodu z firmy je u atributu *Active* nastaveno na *false* - u všech těchto změn dochází pouze k přepsání dané hodnoty.
- *Group* – u této dimenze je nejvíce očekávána změna země, která ale i přesto není příliš častá; ve výjimečných případech je možná změna u názvu skupiny či jeho jednotky; je-li skupina vyřazena, je hodnota atributu *Active* nastavena na *false*.
- *Customer* – je minimálně předpokládána změna u názvu zákazníka, či názvu jeho nadřazené společnosti.
- *Service* – u dimenze služeb se příliš se změnami nepočítá, při vyřazení služby z katalogu služeb je do hodnoty atributu *Lifecycle* zapsáno *Retired*.
- *Configuration Item* – očekávání změn je podobné jako u dimenze služby.
- *Ticket* – u dimenze tiket může před uzavřením samotného tiketu docházet k aktualizaci u atributů *Priority*, *State*, *SLA solving*, *SLA reaction*, *Resolved* a *Closed*. Po uzavření tiketu již žádné změny nenastanou.

## Speciální dimenze

V dimenzionálním modelu nejsou zahrnuty žádné speciální typy dimenzí, jež byly v teoretické části zmíněny. Výjimkou je pouze degenerovaná dimenze, v níž je obsažen jediný atribut zajišťující v tabulce faktů jedinečnost záznamů jako součást složeného primárního klíče. Degenerovaná dimenze je popsána v tabulce Tab. 4.15.

<b>Dimenze:</b>	Deg MI		
<b>ID:</b>	Deg_dim_mi		
<b>Cíl</b>		<b>Zdroj</b>	
<b>Atribut</b>	<b>ID</b>	<b>Tabulka + atribut</b>	<b>Pravidla ETL</b>
ID	ID	-	PK, auto-increment
IDDeg	idDeg	<b>metric_instance{sys_id}</b>	ZP

Tab. 4.15 Degenerovaná dimenze (Zdroj: vlastní zpracování)

#### 4.1.4. Identifikace faktů

Ve fázi identifikace faktů jsou specifikována numerická fakta, jež budou popisována podle výše vytýčených dimenzí. Již v kapitole *Stanovení granularity* byly ke zvolené granularitě předběžně určena fakta – šlo o otevřené, vyřešené, zavřené, propadlé, znovuotevřené tikety, spokojenost uživatelů, průchody tiketů, tikety vyřešené na SD, na první úrovni, při prvním kontaktování, tikety přeposlané na vyšší tier.

V tabulce Tab. 4.15 jsou shrnuta veškerá fakta včetně jejich popisu na určení typu faktu (aditivní, semi-aditivní, neaditivní), která byla stanovena na základě vytýčeného procesu a diskuze s manažery oddělení a konzultantem na KPI.

<b>Fakta:</b>	Day Revue	
<b>ID:</b>	Fact_Day_Revue	
<b>Fakta</b>	<b>Popis</b>	<b>Typ</b>
opened_tickets	Počet otevřených tiketů	aditivní
resolved_tickets	Počet vyřešených tiketů	aditivní
closed_tickets	Počet zavřených tiketů	aditivní
breached_tickets	Počet propadlých tiketů	aditivní
end_user_feedback	Průměrné ohodnocení poskytnuté služby zákazníkovi	semi-aditivní
touches_person	Počet doteků uživatele při práci na tiketech	aditivní
touches_group	Počet doteků připadající skupině při práci uživatelů dané skupiny na tiketech	aditivní
send_to_tier	Počet tiketů přeposlaných SDčkem na vyšší tier	aditivní
only_sd	Počet tiketů vyřešených SDčkem, kteří jako první tiket obdrželi a začali ho řešit	aditivní
first_level	Počet tiketů vyřešených jen na úrovni SD (nebyly přeposlány na vyšší tier)	aditivní

first_contact	Počet tiketů vyřešených při prvním kontaktování SD - u kontaktního typu chat a phone podle příznaku ( <i>only_handled_by_SD=TRUE</i> ) + u kontaktního typu email a integrace je durace řešení do hodiny a tiket zůstal na původně kontaktovaném SD	aditivní
netto_duration	Součet aktivní (čisté) práce specialistů na tiketech; (aktivní stav = <i>Work in progress, Assigned, Resolved - Awaiting Verificatio</i> )	semi-aditivní

Tab. 4.16 Tabulka faktů Day Revue (Zdroj: vlastní zpracování)

Pro dimenzionální model byla vytvořena jediná tabulka faktů, tudíž se v modelu nevyskytují žádné případy speciální tabulek faktů jako sdílená tabulka faktů, bezfaktová tabulka faktů, či událostní tabulka faktů.

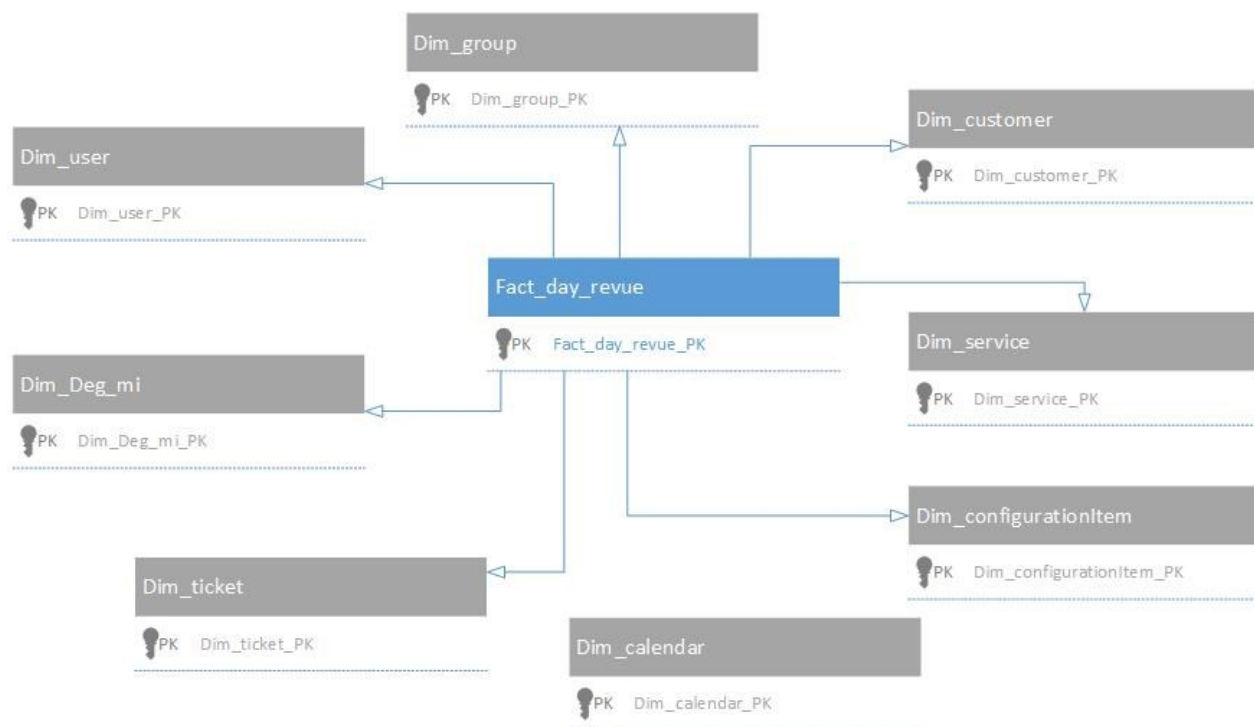
### Primární klíč tabulky faktů

Jak bylo v teoretické části řečeno, primární klíč tabulky faktů je ve většině případů složen z cizích klíčů dimenzí, které jsou s danou tabulkou faktů ve vztahu. Není-li tímto složeným klíčem zajištěna jedinečnost záznamu, je k tabulce faktů připojena tzv. degenerovaná dimenze, kde je jeho atribut přidán k ostatním cizím klíčům, čímž je zaručena unikátnost záznamu.

Složený klíč ve výše zmíněné tabulce faktů je tvořen z jednotlivých umělých klíčů navržených dimenzí, jde o *userID*, *groupID*, *customerID*, *serviceID*, *configurationItemID*, *ticketID*. Tímto primárním klíčem není zabezpečena jedinečnost záznamů, proto musela být vytvořena speciální degenerovaná dimenze, kde je umělý klíč *ID* přidán k ostatním cizím klíčům.

### Konceptuální / logický / fyzický model

V této fázi dimenzionální modelování je vše navrženo pro to, aby mohl být sestaven konceptuální model jako grafické shrnutí tohoto modelu. Konceptuální model je zachycen na obrázku Obr. 4.1.



Obr. 4.1 Konceptuální model (Zdroj: vlastní zpracování)

Rovněž je možné sestavit logický model, v němž jsou již specifikovány konkrétní atributy jednotlivých entit. Model je vyobrazen v příloze 2. Pro vytvoření konceptuálního a logického modelu byl využit software Visio 2013 určený pro tvorbu UML diagramu od společnosti Microsoft. Oba tyto modely byly po konzultaci schváleny a budou v další části práce využity.

#### 4.1.5.Fyzický návrh

Z teoretické části práce je známo, že ve fyzickém návrhu je věnována pozornost agregaci, indexaci a návrhu na rozdělení tabulek. Aplikováním těchto parametrů je možné docílit rychlejší odezvy při vzniku požadavku na databázi.

##### Agregace dat

Předkalkulováním záznamů na požadovanou hladinu agregace může být dosaženo efektivnějšího získání výsledku. V rámci diplomové práce je uživatelům umožněno provádět agregace přímo v nástroji pro reporting a analýzu. Agregace údajů je odlišná od úrovně řízení v oddělení. Žádná agregace ve fyzickém návrhu není uplatněna, neboť nebyl zjištěn optimální stupeň agregace.

## Indexace dat

Smysl indexů byl popsán v kapitole 2.3.7 a v této fázi jsou indexy prakticky realizovány. Oproti agregaci jsou indexy v dimenzionálním modelu implementovány.

Indexy v tabulkách dimenzí jsou standardně umístěny na primárních klíčích jednotlivých dimenzí, jde o indexy- *userID*, *groupID*, *customerID*, *service ID*, *configurationItemID*, *dateID*, *ID*, *ticketID*. Mimo výchozí indexy jsou v následujících dimenzích vytvořeny tyto neunikátní indexy:

- *User* – na attributech *user\_name* (*user*), *email* (*email*), *primary\_group* (*pmGroup*), *manager\_name* (*mngName*), *manager\_email* (*mngEmail*).
- *Group* – na atributu *group\_name* (*group\_name*).
- *Ticket* – na attributech *number* (*number*), *opened* (*opened*), *resolved* (*resolved*), *closed* (*closed*).
- *Calendar* – na attributech *date* (*date*), *month* (*month*), *month\_name* (*month\_name*), *qtr* (*qtr*), *year* (*year*).

U zbylých dimenzí není potřeba speciálních indexů, neboť v nich nebude uloženo velké množství záznamů.

V tabulce faktů *Day Revue* je vytvořen jeden zřetězený index složený z cizích klíčů jednotlivých dimenzí.

## Návrh na rozdělení tabulek

Rozdělení některých z tabulek nebude v části fyzického návrhu tohoto dimenzionálního modelu využito.

Po této fázi byl vyhotoven fyzický model (příloha 3), a to v nástroji pro modelování dat SQL Power Architect. Byly navrženy jednotlivé tabulky dimenzí a faktů společně s odpovídajícími atributy a konkrétními datovými typy a různými omezeními. Tímto nástrojem byl vygenerován SQL skript (příloha 4) pro faktické vytvoření dimenzionálního modelu na vybrané databázové platformě MS SQL Server 2012. Dimenze a tabulka faktů byly vytvořeny a jsou jako relační databáze využity v dalších krocích diplomové práce.

## 4.2. Implementace datových pump

V této praktické části diplomové práce jsou zpracovány informace týkající se nahrávání dat do jednotlivých tabulek dimenzionálního modelu pomocí navržených a nastavených tzv. *jobů*



(rutin). Dle Ballard (2012) jde o integrační a transformační úlohy v BI projektech zajišťující ve své základní podstatě výběr dat z produkčních databází a jejich transformace do databází datového skladu, resp. datového tržiště.

### 4.2.1. Určení zdrojových a cílových datových struktur

V první řadě je nutné specifikovat zdrojové a cílové datové struktury, aby bylo jasné z jakých zdrojů a jaká data budou čerpána a do jakých úložišť a jaká data budou nahrána.

Data do dimenzionálního modelu jsou čerpána z jediného zdroje, a to ze zkušební produkční databáze. Ta je uložena v prostředí MS SQL Server 2012. Jak již bylo výše řečeno, jde o stovky tabulek, z nichž jsou používány jen ty potřebné. Výjimkou je soubor ve formátu .csv, z kterého jsou čerpána data pro dimenzi *Calendar*. Žádný jiný zdroj pro tuto práci využit nebude.

Jako cílové úložiště je určen výše navržený dimenzionální model, jehož implementace byla provedena rovněž v databázi MS SQL Server 2012.

K transformaci dat z jednoho zdroje do druhého je využit *Open source* nástroj pro integraci dat a návrh jobů - *Talend Open Studio for Data Integration*. Nejprve je pomocí průvodce nastaveno spojení do obou databází – zdrojové i cílové (Obr. 4.2) a následně jsou načtena jednotlivá schémata vybraných tabulek.

The screenshot shows the 'Database Properties' dialog box in Talend Open Studio. The 'DB Type' is set to 'Microsoft SQL Server'. The 'String of Connection' field contains 'jdbc:jtds:sqlserver://WL300105:1433/DP\_dimM;instance=SQLEXPRESS'. The 'Login' field is 'sa' and the 'Password' field is masked with dots. The 'Server' field is 'WL300105', the 'Port' is '1433', the 'DataBase' is 'DP\_dimM', and the 'Schema' field is empty. A 'Check' button is located at the bottom right of the connection fields. Below the connection fields, the 'Database Properties' section shows 'SQL Syntax' set to 'SQL 92', 'String Quote' set to 'instance=SQLEXPRESS"', and 'Null Char' set to '000'.

Obr. 4.2 Vytvoření databázového spojení do cílové databáze (Zdroj: vlastní zpracování)

Také je připojen externí soubor, jenž budou sloužit jako kalendář v dimenzionálním modelu a je nahráno jeho schéma.

### 4.2.2.Kontrola dat

Jelikož jsou data extrahována pouze ze zdrojové databáze, která je v podniku k práci používána většinou zaměstnanců, jsou tato data v dostatečné kvalitě. Není potřeba při extrakci věnovat pozornost kvalitě dat.

### 4.2.3.Návrh rutin

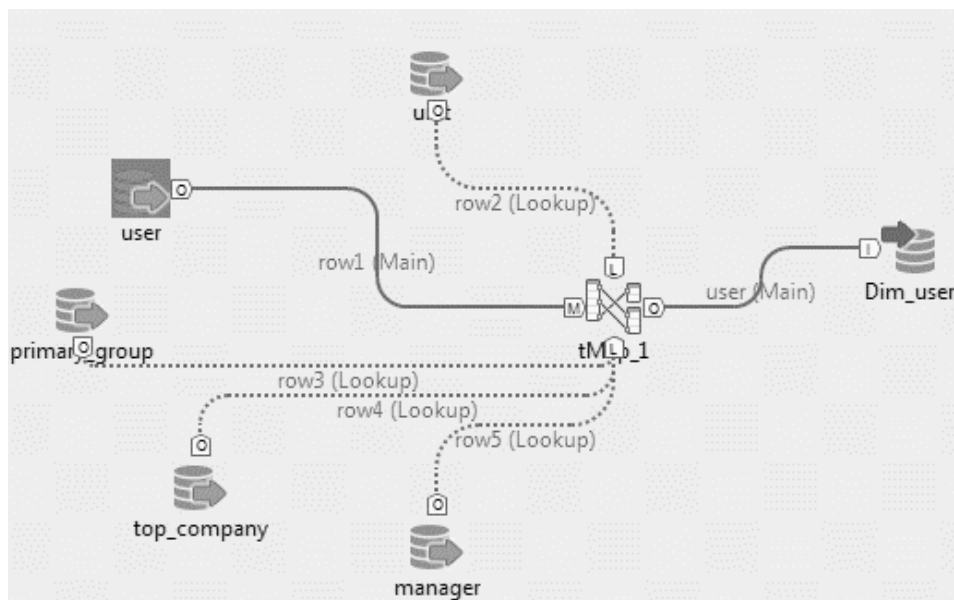
V této fázi jsou navrhovány *rutiny* pro transformaci dat ze zdrojových systémů do cílových struktur. *Rutiny* musí být spouštěny v takovém pořadí, aby při použití dimenzí jako zdrojové tabulky byla tato dimenze již daty naplněna. Např. může jít o tabulku faktů, do které jsou ukládány právě umělé klíče dimenzí.

Pro návrh rutin je využito komponent nabízených vybraným nástrojem. Jelikož jsou obě databáze stejného poskytovatele MS SQL Server, jsou použity jeho komponenty, zejména – *tMSSqlInput* pro zdroj dat a *tMSSqlOutput* pro cílovou databázi. Dále je využito mapovací funkce *tMap*, která je určena pro mapování atributů mezi jednotlivými zdroji a jejich úpravy do předem daných struktur. Před každým nahráním dat do navržených tabulek je použita komponenta *tLogRow*, jež slouží pro případné zjištění chyby při plánování rutin a zkušební zobrazení výsledné dimenze.

Data jsou ze zdrojových tabulek získávána pomocí SQL příkazu *SELECT* a do cílových tabulek jsou v případě historických dat nahrávána příkazem *INSERT*.

#### **Rutina pro dimenzi *User***

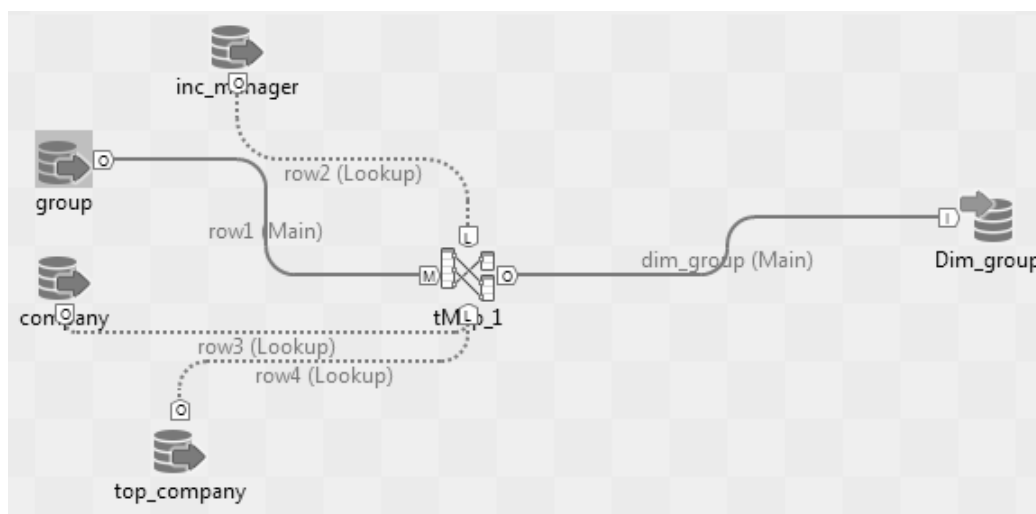
Dimenze *User* je fyzicky složena ze tří zdrojových tabulek *sys\_user*, *sys\_user\_group*, *core\_company*. Ty jsou mezi sebou propojeny a některé reverzně samy na sebe, aby mohly být naplněny všechny navržené atributy dimenze. Nemusely být nastaveny žádné změny datových typů, vše odpovídalo primárnímu zdroji. Návrh rutiny pro tuto dimenzi je na obrázku Obr. 4.3.



Obr. 4.3 Rutina pro dimenzi User (Zdroj: vlastní zpracování)

### Rutina pro dimenzi Group

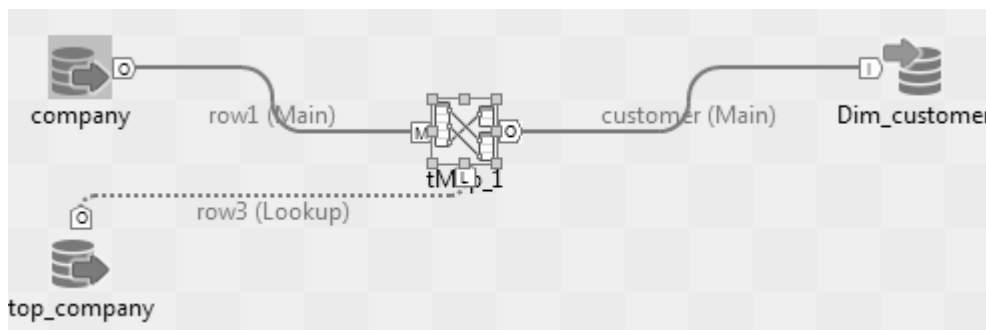
Pro naplnění dimenze Group je zapotřebí tří skutečných tabulek ze zdrojové databáze *sys\_user\_group*, *sys\_user*, *core\_company*, kde je jako hlavní tabulka *sys\_user\_group*. Ty jsou pomocí mapovací funkce mezi sebou spojeny a namapovány do atributů této dimenze (Obr. 4.4). Nebylo potřeba upravovat datové typy atributů.



Obr. 4.4 Rutina pro dimenzi Group (Zdroj: vlastní zpracování)

### Rutina pro dimenzi Customer

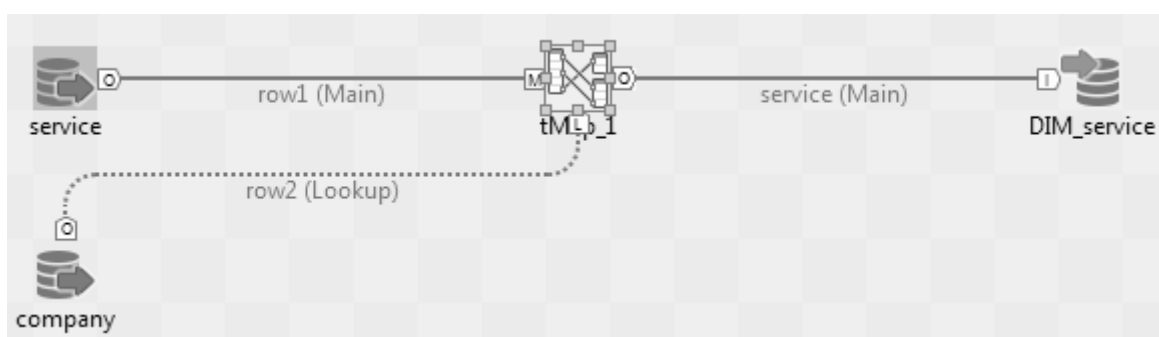
Rutina pro dimenzi Customer je složena pouze z jedné zdrojové tabulky *core\_company*, která se odkazuje na sebe sama. Před nahráváním nebyla nutná žádná změna datových typů.



Obr. 4.5 Rutina pro dimenzi Customer (Zdroj: vlastní zpracování)

### Rutina pro dimenzi Service / Configuration Item

Data do rutin pro dimenze *Service* a *Configuration Item* byla čerpána ze stejných zdrojových tabulek *cmdb\_ci* a *core\_company*, proto je vyhotoven obrázek Obr. 4.6 pouze pro dimenzi *Service*, pro *Configuration Item* je rutina velmi podobná. Není zapotřebí jiných úprav, avšak u dimenze *Service* je nastavena podmínka pro konfigurační položku, která musí být rovna *service\_offering*.

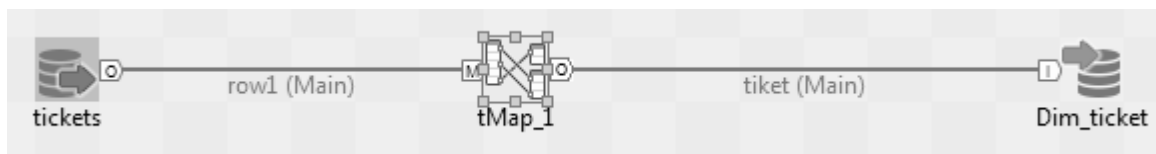


Obr. 4.6 Rutina pro dimenzi Service (Zdroj: vlastní zpracování)

### Rutina pro dimenzi Ticket / degenerovaná dimenze Deg MI

Rutiny pro dimenzi *Ticket* a degenerovanou dimenzi *Deg MI* jsou téměř stejné, proto je zobrazena pouze jedna rutina pro dimenzi *Ticket* na obrázku Obr. 4.7. Data pro dimenzi *Ticket* jsou získána ze zdrojové tabulky *task* a pro degenerovanou dimenzi z tabulky *metric\_instance*. V tabulce *task* je upraven příkaz pro výběr datových hodnot (*opened\_at*, *resolved\_at*, *closed\_at*). Ve zdrojové tabulce jsou hodnoty uchovávány i s časovou jednotkou a ta není v dimenzi *Ticket* potřebná. Pomocí funkce `cast(task.closed_at as date)` byla hodnota transformována pouze na datum. U těchto atributů musely být také ve zdroji pozměněny datové typy na *Date*, neboť byly automaticky nastaveny na *String*. Rovněž musel být pozměněn vzor pro

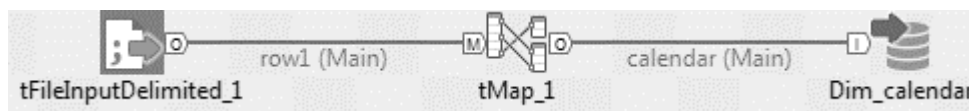
datum (*Data Pattern*) u zdrojové tabulky na „dd-MM-yyyy“ a u cílové na „dd/MM/yyyy“. Degenerovanou dimenzi nebylo nutné nijak upravovat.



Obr. 4.7 Rutina pro dimenzi Ticket (Zdroj: vlastní zpracování)

### Rutina pro dimenzi *Calendar*

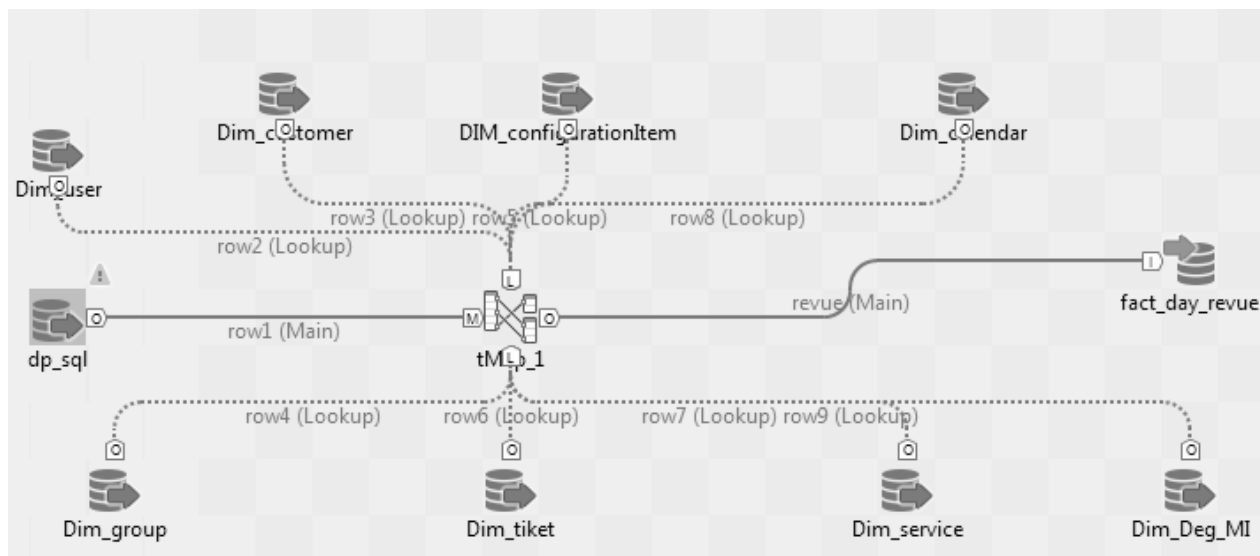
Tato rutina je specifická v tom, že data do dimenze *Calendar* jsou čerpána z jiného zdroje než předešlé rutiny, a to z externího souboru *.csv* za pomoci komponenty *fFileInputDelimited*. V této rutině musely být také upraveny některé datové typy týkající se data, nebo číselných polí. Rovněž byl pozměněn vzor pro datum u zdrojového souboru na „dd.MM.yyyy“ a u cílové tabulky tak, jak v předešlé rutině.



Obr. 4.8 Rutina pro dimenzi Calendar (Zdroj: vlastní zpracování)

### Rutina pro tabulku faktů *Day Revue*

Poslední rutinou je rutina pro tabulku faktů *Day Revue*. Hlavním zdrojem dat je příkaz *SELECT* jako vstupní zdroj nad produkční databází (viz příloha 5). Po vložení příkazu do vstupní komponenty musí být odhadnuto schéma příkazu, aby bylo možné provést namapování s dimenzemi a výslednou tabulkou faktů. Dále jsou k mapovací komponentě připojeny vytvořené a naplněné dimenze (Obr. 4.9). Pomocí mapovací komponenty dojde k propojení odpovídajících atributů příkazu s produkčními klíči dimenzí. Nakonec jsou umělé klíče dimenzí vloženy do odpovídající struktury navržené tabulky faktů rovněž s fakty souhrnného SQL příkazu.



Obr. 4.9 Rutina pro tabulku faktů Day Revue (Zdroj: vlastní zpracování)

#### 4.2.4. Základní transformace

Po spuštění rutin jednotlivých dimenzí dojde k jejich naplnění daty. V tomto modelu nezáleží na pořadí spuštěných rutin dimenzí, neboť dimenze nejsou na sobě závislé, avšak musí být spuštěny před nahráváním dat do tabulky faktů. Po nahrání dat do dimenzí je spuštěna rutina pro tabulku faktů. Při nahrávání rutin je dostačující využití základní transformace dat.

### 4.3. Návrh výstupů a jejich prezentace

Cílem poslední kapitoly diplomové práce je s využitím *User Experience*, o kterém bylo hovořeno v teoretické části, navrhnout výstupní dashboardy pro koncové uživatele. V první řadě jsou sesbírány požadavky od uživatelů a je sestaven hrubý návrh výstupů. V konečné fázi jsou dashboardy fyzicky navrženy v nástroji Qlik Sense.

#### 4.3.1. Sběr požadavků výstupů

Mnoho požadovaných informací, které uživatelé potřebují sledovat, bylo zmíněno v kapitolách 3.4.1 a 4.1.1 - na všechny je v této části navázáno. Zde je zjišťováno jaké informace by dashboardy měly poskytovat, jakou formou by měly být prezentovány a jaké jsou speciální požadavky uživatelů na výstupy.

Pro každou vytyčenou úroveň manažerů by měl být sestaven výchozí dashboard, kterým jsou poskytovány základní a shrnující informace za celou zkoumanou oblast. V následujících

dashboardech je možné se zabývat detailnějšími analýzami. Byly sepsány následující požadavky dle jednotlivých úrovní manažerů:

- **Manažeři vyšších úrovní**

Jak bylo v kapitole 3.4.1 zmíněno, je pro tyto manažery důležité sledovat výsledky za celé oddělení převážně na měsíční bázi. Na výchozím dashboardu by měly být zobrazeny informace týkající se metrik *First Contact Resolution*, *First Level Resolution*, *Keeping SLA* v porovnání s jejich optimální hodnotou. Dále by měl být v ročním trendu zachycen měsíční vývoj tiketů *First Contact*, *First Level*, *Breached tickets*. Mělo by být umožněno filtrování dle let a měsíců a také vybrání konkrétní země.

Detailnější dashboardy by měly být sestaveny dle logického uvážení. Jde o informace týkající se vývoje otevřených / vyřešených / zavřených tiketů hlavně k prioritě tiketů, trend průchodů tiketů, úrovni řešení, úrovni znovu otevření.

- **Manažeři středních úrovní**

Pro manažery střední úrovně je typické monitorovat měsíční, příp. týdenní informace o jednotlivých skupinách SD v jeho vedené zemi. Na dashboardech by měla být možnost filtrování výsledků dle let, měsíců, týdnů, zemí a skupin. Na výchozím dashboardu by měly být informace týkající se otevřených / vyřešených / zavřených tiketů vzhledem k výše zmíněným filtrům, ale také k prioritě a kontaktnímu typu spolu s úrovní řešení.

Konkrétně zaměřenými dashboardy by měly být zachyceny informace týkající se výše zmíněných ukazatelů se zaměřením na zákazníka, službu, konfigurační položku. Dále by měly být zachyceny informace propadlých tiketů se zaměřením na prioritu, tiketů poslaných na vyšší tier, tiketů vyřešených na úrovni daného SD, znovu otevřených tiketů.

- **Manažeři nižších úrovní**

Manažeři nižších úrovní potřebují kontrolu nad službou na denní úrovni. Pro tyto manažery by na základním dashboardu měly být viditelné údaje o odpovídající SD skupině a SD specialitech. Mělo by jít o informace zaměřené na vývoj otevřených / vyřešených / propadlých tiketů. Rovněž vývoj průchodů jednotlivých lidí a průměrný čas řešení tiketu.

V detailnějších dashboardech by se výše zmíněné informace měly vyskytovat s detailem na prioritu tiketů, kontaktní typ, zákazníky a službu.

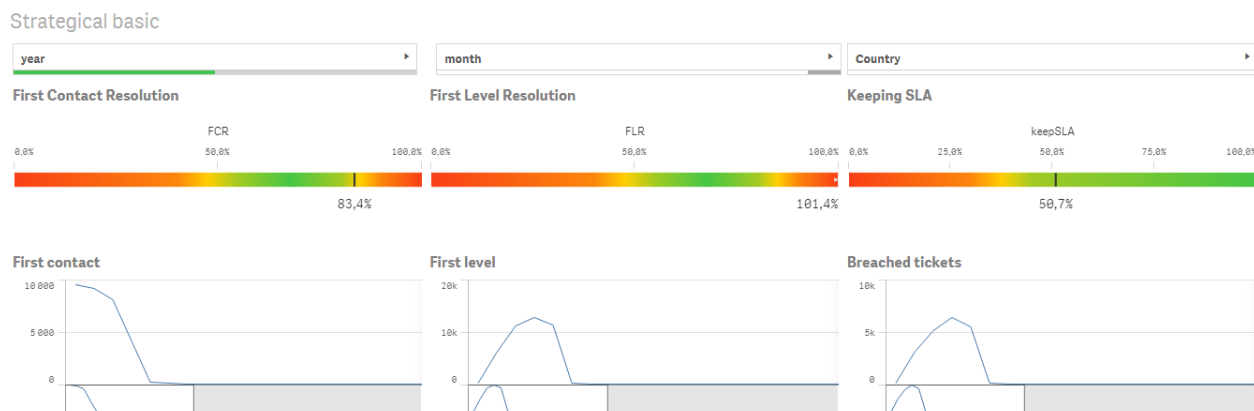
### 4.3.2.Návrh výstupů

Po domluvě s daným oddělením bylo dohodnuto, že pro návrh i prezentaci dashboardů bude využito softwarového nástroje pro vizualizaci dat Qlik Sense, který je již na jiném oddělení k dispozici a je možné s tímto nástrojem do budoucna počítat. Dashboardy pro diplomovou práci jsou vytvořeny na lokálním počítači v bezplatné verzi tohoto software.

Návrhy dashboardů jsou rozděleny do tří úrovní, a to na již řečené úrovní manažerů. Každý dashboard má svou úvodní stránku a výchozí nastavení, které se při načtení dashboardů zobrazí. Pro podrobnější zkoumání je možné jednoduše procházet mezi jednotlivými dashboardy dané úrovně.

Na následujících obrázcích Obr 4.10 – 4.12 je zobrazena část vytvořených dashboardů – pro každou úroveň jeden výstup. Jak již bylo řečeno, k diplomové práci je využita zkušební množina dat, proto informace na dashboardech nekorespondují se skutečnou realitou. Veškeré konkrétní názvy podniku jako jména uživatelů, skupin, zákazníků, služeb, konfigurační položek jsou z bezpečnostních důvodů nedostupná.

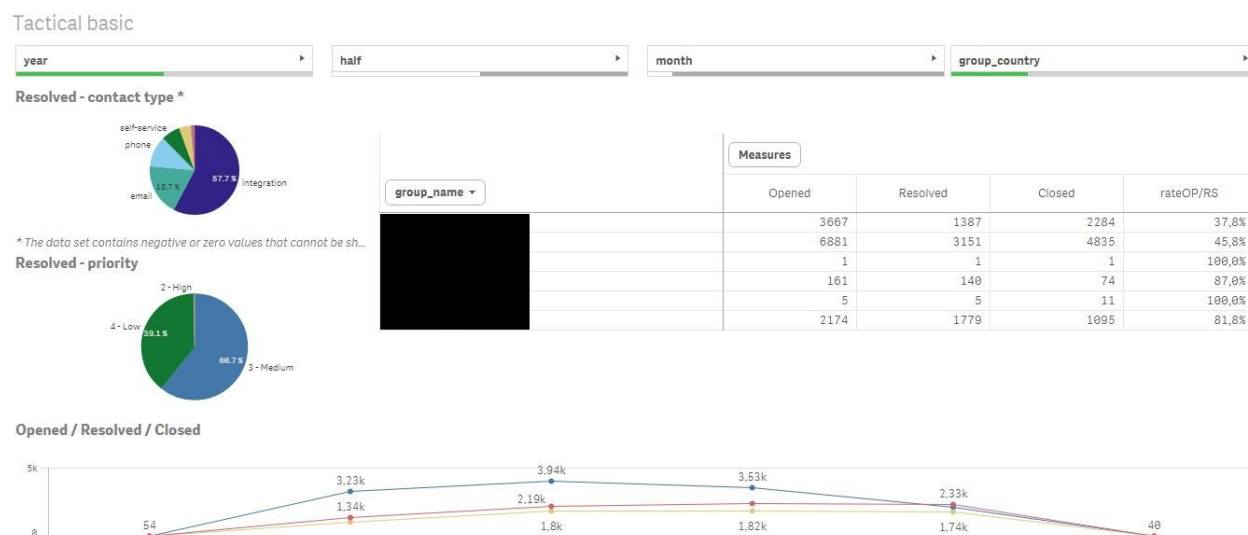
Na obrázku Obr. 4.10 je dashboard pro vyšší úroveň manažerů. Jde o výchozí dashboard poskytující pomocí sloupcových indikátorů základní informace o dosažené hodnotě *First Contact Resolution*, *First Level Resolution* a *Keeping SLA*, kdy má každý ukazatel nastaveno barevné rozlišení / rozmezí pro vyhovující, méně vyhovující a nevyhovující výsledek. Dále je pod těmito ukazateli roční trend měsíčních hodnot – *First Contact*, *First Level* a *breached tickets*. V dashboardu je implementován panel pro filtrování roku, měsíce a země oddělení.



Obr. 4.10 Dashboard pro vyšší úroveň manažerů (Zdroj: vlastní zpracování)

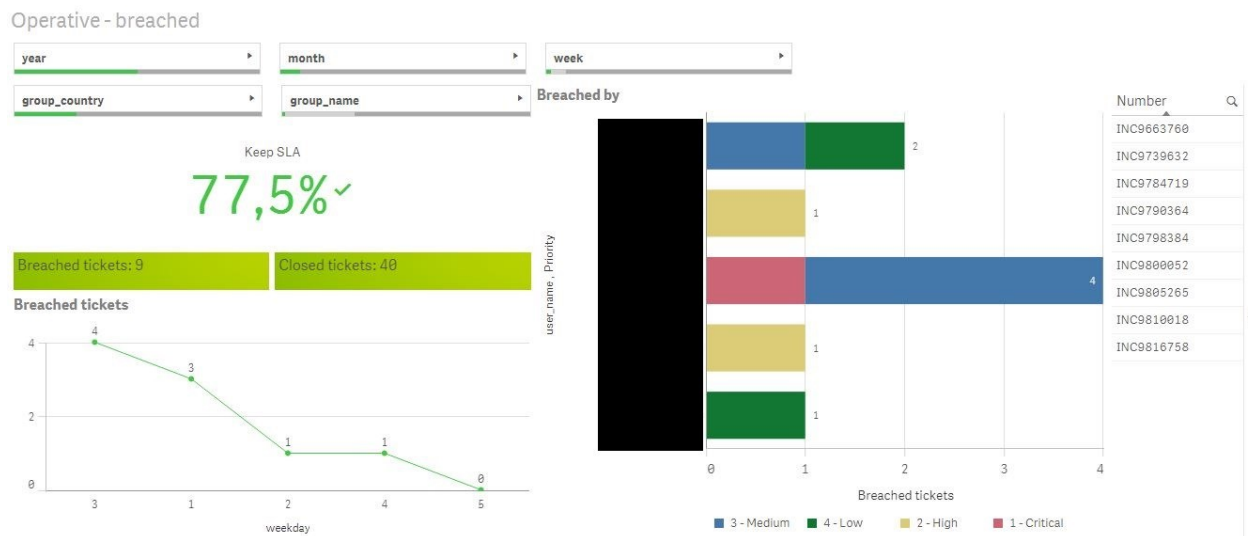


Na obrázku Obr. 4.11 je zobrazen výchozí dashboard pro střední úroveň manažerů. Dashboardem jsou zachyceny základní informace o týdenním vývoji otevřených / vyřešených / zavřených tiketů v ročním trendu. Koláčovými grafy je vyčísleno procento vyřešených tiketů vzhledem k prioritě a kontaktnímu typu. Tabulka poskytuje rychlý přehled o číselných údajích tiketů spolu s úrovní řešení. Jednotlivé části těchto grafů i řádky tabulky jsou určeny k vyfiltrování požadované oblasti. K vybrání požadovaného roku, půlroku, měsíce či země SD je vložen filtrovací panel.



Obr. 4.11 Dashboard pro střední úroveň manažerů

Poslední z ukázkových dashboardů na obrázku Obr. 4.12 je určen pro manažery nižších úrovní, kteří potřebují mít poskytovanou službu denně pod kontrolou. Tento dashboard se od ostatních dvou liší v tom, že nejde o výchozí dashboard poskytující základní informace, ale jde o konkrétně zaměřený dashboard, a to na propadlé tikety. Na dashboardu je možné sledovat týdenní vývoj propadlých tiketů ve dnech. Dále je pomocí sloupcového grafu vyčíslen počet propadlých tiketů na konkrétní SD specialisty vzhledem k prioritě tiketů. Na levé straně dashboardu je umístěna tabulka se zainteresovanými čísly tiketů k nastavené podmínce. Rovněž je vypočteno udržení SLA z počtu propadlých a zavřených tiketů. Ve filtrovacím panelu je možné filtrovat podle roku, měsíce, týdne v roce, země SD a skupiny SD.



Obr. 4.12 Dashboard pro nižší úroveň manažerů

Podobně jsou koncipovány ostatní navržené dashboardy.

### 4.3.3. Zhodnocení výstupů

V této kapitole je provedeno resumé požadavků z Tab. 4.4, zda a jak byly vytvořeny. S oslovenými uživateli byly některé dashboardy otestovány. Pro vytvoření hodnocení navržených dashboardů na základě UX byla sjednána schůzka dvaceti uživatelů, kterými byly dashboardy testovány pro vybrané rámcové požadavky uvedené v Tab. 4.4, a to z následujících pohledů:

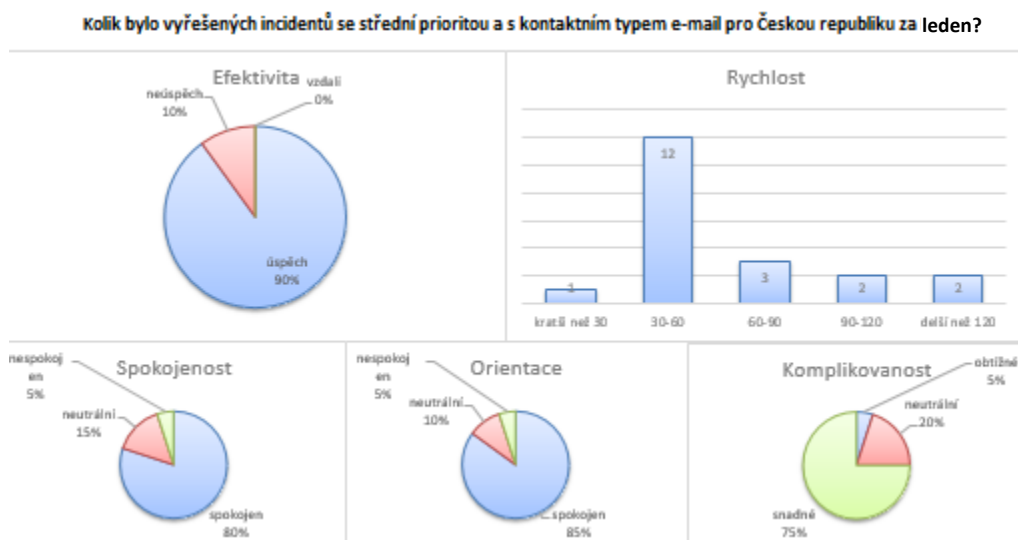
- „*Jaká je efektivita designu?*“ → Zde byli uživatelé testováni na to, zda konečný výsledek našli úspěšně, neúspěšně, anebo hledání vzdali.
- „*Jaká je rychlost hledání?*“ → Jde o zjištění, kolik času uživatelé strávili při hledání zmíněného požadavku. Následně byli rozřazeni do skupin podle toho, zda zapadali do určených intervalů (kratší než 30 s, 30-60 s, 60-90 s, 90-120 s, delší než 120 s).
- „*Jaká je spokojenost s navrhnutým dashboardem?*“ → Uživatelé zde odpovídali, zda jsou s designem spokojeni, nespokojeni, nebo se jim design jeví jako neutrální.
- „*Jaká je v dashboardech orientace?*“ → Zde uživatelé hodnotili orientaci v dashboardech, zda jsou s ní spokojeni, nespokojeni nebo neutrální.
- „*Jak je nalezení výsledku komplikované?*“ → Uživatelé se zde snažili vystihnout komplikovanost nalezení výsledku, zda je snadné, neutrální nebo obtížné.

Provedení a zhodnocení jednotlivých požadavků je následující:

### Požadavek č. 1: Počet otevřených / vyřešených / zavřených incidentů

Tento požadavek je zahrnut do všech tří úrovní dashboardů, které byly v předešlé kapitole stanoveny. Každé úrovni dashboardu je zviditelněna různá úroveň detailu podle toho, co uživatelé potřebují sledovat (např. pro vyšší manažery je vidět celkový počet otevřených / vyřešených / zavřených incidentů s možností drill-down na konkrétní země). Tím je umožněno sledování těchto ukazatelů na skupiny a uživatele, kde jsou tyto výsledky zaznamenány pomocí tabulky. Dále je vývoj těchto hodnot sledován v liniovém grafu. S využitím koláčového grafu je zachyceno procento otevřených / vyřešených / zavřených incidentů vzhledem k prioritě a kontaktnímu typu (ukázka je na Obr. 4.11 na vyřešených incidentech). Dále je pro sledování těchto hodnot vytvořen indikátor, ve kterém je zachycen aktuální stav, ale také optimální hodnota. Ke sledování otevřených / vyřešených / zavřených incidentů z pohledu zákazníka, služby a konfigurační položky je využito filtračního panelu, neboť je těchto hodnot velké množství. Drill-down je zaveden také nad datem a nad zákazníkem.

U toho požadavku bylo testováno – „Kolik bylo vyřešených incidentů se střední prioritou a s kontaktním typem e-mail pro Českou republiku za leden?“ Výsledky z testování uživatelů na položený dotaz je možné vidět na Obr. 4.13.



Obr. 4.13 Hodnocení dashboardu u požadavku č. 1 (Zdroj: vlastní zpracování)

### Požadavek č. 2: Poměr vyřešených incidentů k otevřeným

Tento ukazatel je zachycen pouze v tabulce, a to na všech úrovních dashboardů. Příklad je uveden na Obr. 4.11, kde je poměr vyřešených incidentů k otevřeným uveden pro SD skupiny.

V tabulce je zavedena funkcionální drill-down – země, SD skupina, SD specialista a také drill-down v čase.

### Požadavek č. 3 a č. 4: Počet propadlých incidentů a udržení SLA

Oba tyto požadavky jsou zakomponovány do všech tří úrovní dashboardů, kdy jde o vyjádření počtu incidentů, které propadly a o sledování procenta udržení SLA, což je počet nepropadlých incidentů k uzavřeným. Podle potřeby je každé úrovni umožněna funkcionální drill-down (země, SD skupina, SD specialista a také nad datem). Počet propadlých incidentů je jako trend zaznamenán v liniovém grafu, kde je vidět jejich vývoj v čase (např. Obr. 4.10 a Obr. 4.12). Dále je také využito sloupcového grafu, kde jsou zachyceny propadlé incidenty vzhledem k prioritě. Ke sledování udržení SLA je použito různých indikátorů, kterými je zobrazena aktuální hodnota spolu s barevným odlišením intervalů dobrých, normálních a nevyhovujících hodnot (ukázka na Obr. 4.10 a Obr. 4.12). K filtrování skupin a zákazníků je určen filtrovací panel a k uživatelům sloupcový graf.

U požadavku č. 4 bylo testováno – „*Jaké je procento udržení SLA za celé SD v měsíci lednu?*“ Výsledky z testování uživatelů na položený dotaz je možné vidět na Obr. 4.14.



Obr. 4.14 Hodnocení dashboardu u požadavku č.4 (Zdroj: vlastní zpracování)

### Požadavek č. 5: Spokojenost zákazníků

Spokojenost zákazníků je uvedena v dashboardech pro vyšší a střední manažery, protože tato hodnota je vypočítávána jen pro celé SD, země a skupiny. K zobrazení této hodnoty je využito

indikátorů, kterým je zachycena daná hodnota, a je určeno, do jakého intervalu hodnota zapadá (hodnota mezi 3-4 výborné hodnocení – zelená barva, 2-3 dobré hodnocení – oranžová barva, 1-2 špatné hodnocení – červená barva). Dále je sledován vztah této hodnoty s počtem tiketů vyřešených při prvním kontaktování a také s počtem znovuootevřených incidentů. Uživatelé mohou využít drill-down (země, SD skupina a také nad datem).

Zde měli uživatelé najít hodnotu EUF v měsíci lednu pro vybranou SD skupinu. Výsledky těchto testů byly následující: efektivita – úspěch (95 %), rychlost 30-60 s (16 uživatelů), spokojenost – spokojen (95 %), orientace – spokojen (90 %), komplikovanost – snadné (95 %).

**Požadavek č. 6 a č. 7:** Počet znovuootevřených incidentů a úroveň znovuootevření

Tyto požadavky jsou opět zavedeny na všech třech úrovních dashboardů. Koncept dashboardů těchto požadavků je velmi podobný jako dashboardy u požadavků č. 3 a 4. Počet znovuootevřených incidentů, jak bylo u požadavku č. 5 řečeno, je sledován ve vztahu se spokojeností zákazníka. Na dashboardech je podle potřeby zaveden drill-down (podobný jako v předešlých požadavcích).

**Požadavek č. 8 a č. 9:** Počet doteků a úroveň řešení.

Hodnoty těchto požadavků jsou zavedeny do všech úrovní dashboardů. Počet doteků je sledován v tabulce pro jednotlivé specialisty a skupiny a je sledován rovněž vývoj v čase v liniovém grafu. K pozorování úrovně řešení je využito grafického indikátoru s aktuální hodnotou a barevně odlišenými intervaly, jak bylo použito u většiny výše zmíněných úrovní. Úroveň je také sledována v trendu. Rozdíl je na nejvyšší úrovni dashboardů, tam jsou tyto hodnoty zaznamenány jako průměrná hodnota, a to ze dvou pohledů – průměr za doteky SD specialistů a průměr za doteky SD skupin. Uživatelé mohou využít funkcionalitu drill-down nad skupinou a specialistou.

Zde měli uživatelé najít počet doteků v měsíci lednu pro vybraného SD specialistu. Výsledky těchto testů byly následující: efektivita – úspěch (80 %), rychlost 30-60 s (12 uživatelů), spokojenost – spokojen (95 %), orientace – spokojen (85 %), komplikovanost – snadné (85 %).

**Požadavek č. 10:** Počet přeposlaných incidentů / úroveň přeposlání

Hodnoty tohoto požadavku jsou zahrnuty v dashboardech pro vyšší a střední manažery. Funkcionalita a prvky na dashboardech jsou podobné jako např. u požadavků č. 7 a č. 8. Pomocí

koláčového grafu je navíc zachyceno procentní zastoupení přeposlaných incidentů vzhledem k jejich prioritě.

**Požadavek č. 11:** Průměrný čas strávený na tiketu

V dashboardech na všech třech úrovních je sledován průměrný čas strávený na tiketu v sekundách, a to jak aktuální hodnota, tak i vývoj v čase. S využitím grafických indikátorů je sledován průměrný čas vzhledem k prioritě tiketů, ovšem správně nastavené intervaly času budou muset být ještě vypořizovány. Uživatelé mohou využít funkcionalitu drill-down nad skupinou a specialistou.

**Požadavek č. 12 a č. 13:** Počet tiketů vyřešených na první úrovni / FLR a počet tiketů vyřešených při prvním kontaktování / FCR

Hodnoty pro tyto požadavky jsou uvedeny pouze v dashboardu pro vyšší manažery. U počtu tiketů jde o vývoj hodnoty v čase. U FLR a FCR je použit grafický indikátor s vypočtenou hodnotou, nastavenými intervaly a barevným odlišením.

Uživatelé měli nalézt hodnotu FCR v měsíci lednu. Výsledky těchto testů byly následující: efektivita – úspěch (100 %), rychlost 30-60 s (17 uživatelů), spokojenost – spokojen (95 %), orientace – spokojen (100 %), komplikovanost – snadné (100 %).

Z výše popsaného textu je zřejmé, že byly splněny všechny stanovené požadavky z Tab. 4.4 a do dashboardů byly zakomponovány. Z provedených testů na pět položených dotazů bylo zjištěno, že dashboardy plní svůj účel. Požadované informace jsou rychle k nalezení a není obtížné je nalézt, orientace v nich je intuitivní. Dashboardy byly po této fázi schváleny. Rovněž je ale očekávána další zpětná vazba od ostatních uživatelů a postupné doladění prezentovaných výsledků.

#### **4.3.4. Zveřejnění výstupů**

Tím že jsou výstupy zpracovány v nástroji Qlik Sense a v podniku je tento nástroj na určitých odděleních taky používán, budou dashboardy umístěny na server a budou dostupné z jednoho místa všem, kterým jsou přiřazena práva. Proces přidělování práv není součástí diplomové práce.

## 5. Závěr

Cílem diplomové práce bylo pro oddělení *Customer Service Centre* navrhnout reportingovou platformu pro zpracovávání reportů, dashboardů či jiných zobrazovacích technik, aby všem zainteresovaným uživatelům byly poskytovány stejné a standardizované výsledky oddělení. Důvodem pro vypracování platformy byly nestandardizované metriky a reporty, které mají řadu omezení, velkou možnost vzniku chyb při zpracování a neefektivní prezentaci.

Řešení bylo vyvíjeno na principech dimenzionálního modelování, kde bylo postupováno podle kroků životního cyklu R. Kimballa.

Diplomová práce byla věnována poskytované službě Service Desk, v níž byly nejprve analyzovány procesy. Jako hlavní proces byl vymezen Incident Management a jako podpůrný Service Management zabývající se správou incidentů. Pomocí sémantického modelu byly k těmto procesům sepsány předběžné entity a metriky. Po konzultaci s manažery a konzultantem na KPI byly sepsány veškeré požadavky na výstup, u kterých byla nastavena priorita. Dále byly identifikovány zdrojové systémy pro tvorbu reportů, kdy pro tento model byla jako dostačující zvolena produkční databáze.

V dalších třech krocích následovalo stanovení granularity tabulky faktů, identifikace dimenzí a faktů. Nejprve byla pro vymezené procesy určena granularita, nebo-li úroveň podrobností, s kterou jsou data do tabulky faktů ukládána. Pro tuto granularitu byl určen její typ a předběžné dimenze a fakta. V druhém kroku byly identifikovány dimenze včetně jejich atributů, vazbou na zdrojovou databázi a pravidla pro ETL. Rovněž byla deklarována granularita jednotlivých dimenzí a možné hierarchie. Pozornost byla věnována změnám v dimenzích a speciálním dimenzím. V posledním kroku jsou specifikována numerická fakta. Na základě návrhu byl sestrojen konceptuální a logický model.

Ve fyzickém návrhu byla řešena agregace, indexování a rozdělování tabulek, po kterém byl zkonstruován fyzický model.

Nahrávání dat do vytvořeného dimenzionálního modelu, bylo řešeno implementováním datových pump, kdy nejdříve byly popsány zdrojové a cílové struktury, poté byly navrženy rutiny datových pump pro jednotlivé dimenze a faktovou tabulku.

V poslední části byly s využitím User Experience vytvořeny výstupní dashboardy pro koncové uživatele. Byly sesbírány požadavky na výsledné dashboardy a na jejich základě byly v nástroji Qlik Sense vyhotoveny. Z vyhodnocení požadavků v Tab 4.4 bylo prokázáno splnění všech požadavků. Testováním uživatelů na položené dotazy, bylo zjištěno, že dashboardy plní svůj účel, informace jsou rychle a jednoduše nalezeny a orientace v nich je intuitivní.

Největším přínosem diplomové práce je pro vymezený proces nastavení standardizovaných metrik a KPI, procesu jejich vytváření a dále zajištění uceleného místa a zdrojů dat pro vytváření reportů či dashboardů. Další výhodou je možnost interaktivního reportingu v podobě poskytování výsledků různými grafickými objekty se schopností porovnávání a tím vytvoření lepších podmínek pro rozhodování o budoucích krocích v oddělení.

V rámci diplomové práce se podařilo navrhnout a implementovat reportingovou platformu, kdy při vytváření došlo ke standardizování KPI a metrik a k ucelení jednotného řešení, čímž je zvýšená vypovídací schopnost dat, lepší kontrola nad poskytovanou službou a v oddělení jsou rovněž zefektivněny procesy vzhledem ke kontinuálnímu zlepšování. Výsledné dashboardy budou na vybrané skupině uživatelů odzkoušeny, budou doladěny případně nalezené nedostatky a poté se může uskutečnit nasazení platformy do produkce.

Do budoucna se očekává neustále udržování, zlepšování a rozšiřování modelu. Může jít o důkladnější zaměření na kvalitu dat, na připojení dalších oddělení, přidání dalších standardních metrik a KPI, či dodání nových interních nebo externích zdrojů dat.



## Seznam použité literatury

### Knižní výtisky

KIMBALL, Ralph a Margy ROSS. *The data warehouse toolkit: the definitive guide to dimensional modeling*. 3rd ed. Indianapolis, Ind.: Wiley, c2013.

TVRDÍKOVÁ, Milena. *Aplikace moderních informačních technologií v řízení firmy: nástroje ke zvyšování kvality informačních systémů*. Praha: Grada, 2008. ISBN 978-80-247-2728-8.

POUR, Jan, Miloš MARYŠKA a Ota NOVOTNÝ. *Business intelligence v podnikové praxi*. Praha: Professional Publishing, 2012. ISBN 978-80-7431-065-2.

NĚMEC, Radek. *Principy projektování a implementace systémů Business Intelligence*. Ostrava: VŠB-TU Ostrava, 2014. ISBN 978-80-248-3452-8.

KALUŽA, Jindřich. *Informační systémy pro strategické řízení*. Ostrava: Vysoká škola báňská - Technická univerzita, Fakulta ekonomická, c2010. ISBN 978-80-248-2280-8.

MUNDY, Joy, Warren THORNTHWAITE a Ralph KIMBALL. *The Microsoft data warehouse toolkit: with SQL server 2008 R2 and the Microsoft Business Intelligence toolset*. 2nd ed. Indianapolis, IN: Wiley, c2011.

BALLARD, Chuck. *Dimensional modeling: in a business intelligence environment*. San Jose, Calif.: IBM International Technical Support Organization, 2006. ISBN 0738496448.

GÁLA, Libor, Jan POUR a Zuzana ŠEDIVÁ. *Podniková informatika: počítačové aplikace v podnikové a mezipodnikové praxi*. 3. aktualizované vydání. Praha: Grada Publishing, 2015. ISBN 978-80-247-5457-4.

LACKO, Ľuboslav. *Databáze: datové sklady, OLAP a dolování dat s příklady v Microsoft SQL Serveru a Oracle*. Brno: Computer Press, 2003. ISBN 80-7226-969-0.

HAN, Jiawei a Micheline KAMBER. *Data Mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers, 2001. ISBN 1-55860-489-8.

ECKERSON, Wayne W. *Deploying Dashboards and Scorecards*. United States: The Data Warehousing Institute, 2006.

CARTLIDGE, Alison a spol. *An Introductory Overview of ITIL V3*. The United Kingdom: itSMF, 2007. ISBN 0-9551245-8-1.

### **Elektronické zdroje**

GARTNER. *Business Intelligence* [online] 2015 [cit. 11.12.2015]. Dostupné z: <http://www.gartner.com/it-glossary/business-intelligence-bi>

LEARNDWBI a. Design Star Schema [online] 19.07.2015a [cit. 8.2.2016]. Dostupné z: <http://learndatamodeling.com/blog/designing-star-schema/>

LEARNDWBI. Design Snowflake Schema [online] 25.07.2015b [cit. 8.2.2016]. Dostupné z: <http://learndatamodeling.com/blog/designing-snowflake-schema/>

ZAĀANE, Osmar R. *Database Management Systems: Data Warehousing* [online] 2004 [cit. 9.2.2016]. Dostupné z: <http://webdocs.cs.ualberta.ca/~zaiane/courses/cmp391/slides/L5-391-04.pdf>

GUBE, Jacob. *What Is User Experience Design? Overview, Tools And Resources* [online] 5.10.2010 [cit. 10.2.2016]. Dostupné z: <https://www.smashingmagazine.com/2010/10/what-is-user-experience-design-overview-tools-and-resources/#ux-design-is-not-one-size-fits-all>

GIBSON, Nick. *Dashboard Design: 10 Tips to Design User-Friendly Dashboards* [online] 7.5.2014 [cit. 10.2.2016]. Dostupné z: <https://blog.udemy.com/dashboard-design/>

TOBOLKA, Jiří. *Stavíme efektivní dashboard* [online] 2013 [cit. 10.2.2016]. Dostupné z: <http://m.systemonline.cz/business-intelligence/stavime-efektivni-dashboard.htm>

## **Seznam zkratek**

BI – Business Intelligence  
ETL – Extract, Transform, Load  
ELT – Extract, Load, Transform  
EAI – Enterprise Application Integration  
DSA – Data Staging Area  
DW – Data Warehouse  
DM – Data Mart  
ODS – Operation Data Store  
OLAP – OnLine Analytical Processing  
UML – Unified Modeling Language  
KPI – Key Performance Indicators  
SCD – Slowly Changing Dimensions  
RCD – Rapidly Changing Dimensions  
SQL – Standard Query Language  
BSC – Balance Scorecard  
CSC – Customer Service Centre  
SA – Service Areas  
ITIL – Information Technology Infrastructure Library  
IT – Information Technology  
GSOC – Global Service Operations Centre  
SD – Service Desk  
ITSM – Information Technology Service Management  
SPOC – Single Point of Contact  
SM – Service Management  
PM – Problem Management  
CD – Control Desk  
SLA – Service Level Agreement

EUUF – End User Feedback

VBA – Visual Basic for Application

FLR – First Level Resolution

FCR – First Contact Resolution

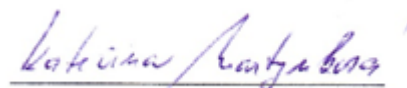
SW - Software

# Prohlášení o využití výsledků diplomové práce

Prohlašuji, že

- jsem byla seznámena s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. – autorský zákon, zejména § 35 – užití díla v rámci občanských a náboženských obřadů, v rámci školních představení a užití díla školního a § 60 – školní dílo;
- beru na vědomí, že Vysoká škola báňská – Technická univerzita Ostrava (dále jen VŠB-TUO) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užít (§ 35 odst. 3);
- souhlasím s tím, že diplomová práce bude v elektronické podobě archivována v Ústřední knihovně VŠB-TUO a jeden výtisk bude uložen u vedoucího diplomové práce. Souhlasím s tím, že bibliografické údaje o diplomové práci budou zveřejněny v informačním systému VŠB-TUO;
- bylo sjednáno, že s VŠB-TUO, v případě zájmu z její strany, uzavřu licenční smlouvu s oprávněním užít dílo v rozsahu § 12 odst. 4 autorského zákona;
- bylo sjednáno, že užít své dílo, diplomovou práci, nebo poskytnout licenci k jejímu využití mohu jen se souhlasem VŠB-TUO, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly VŠB-TUO na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Ostravě dne 22. dubna 2016



Bc. Kateřina Martynková

# **Seznam příloh**

Příloha 1: Incident Management Process

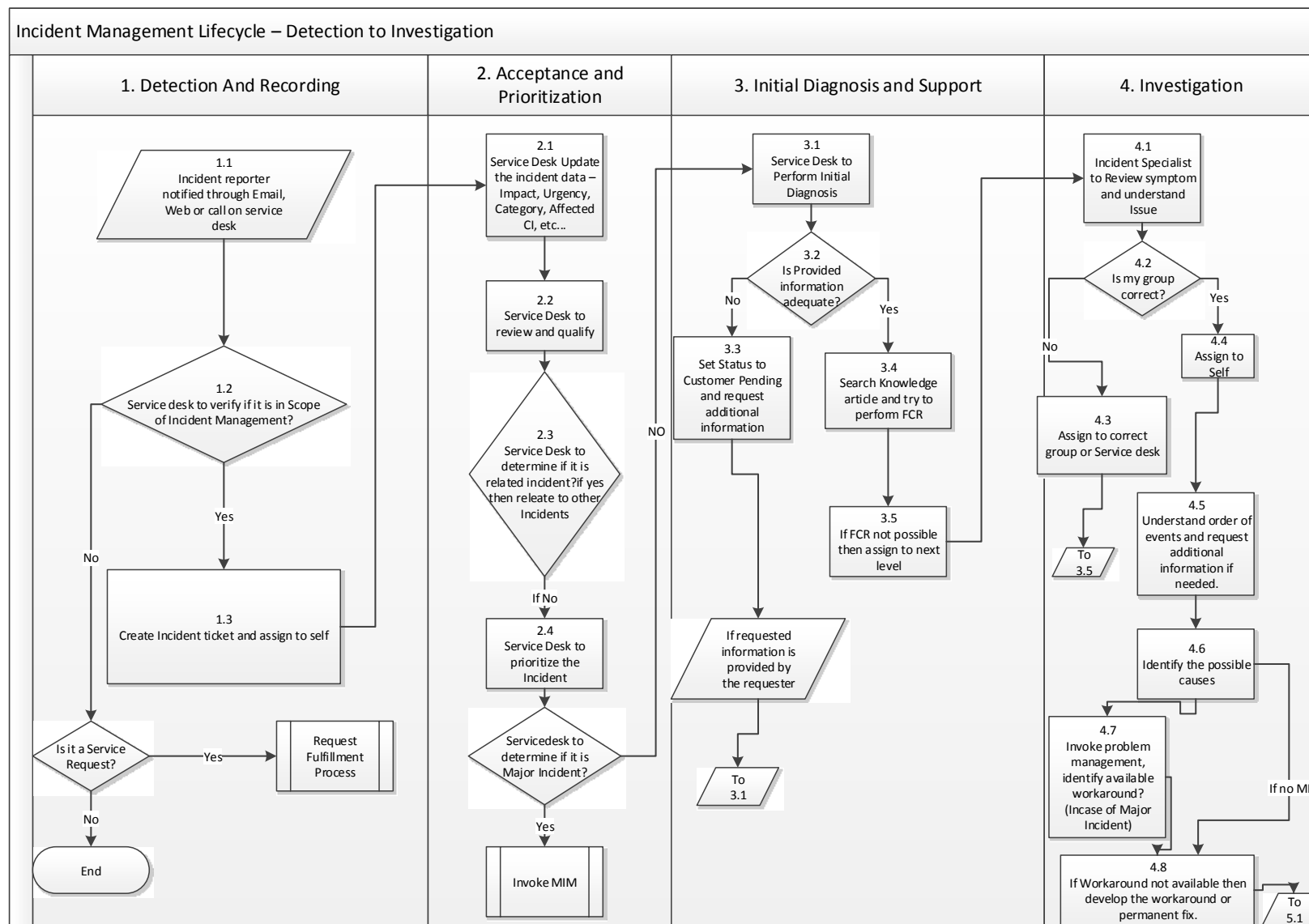
Příloha 2: Logický model

Příloha 3: Fyzický model

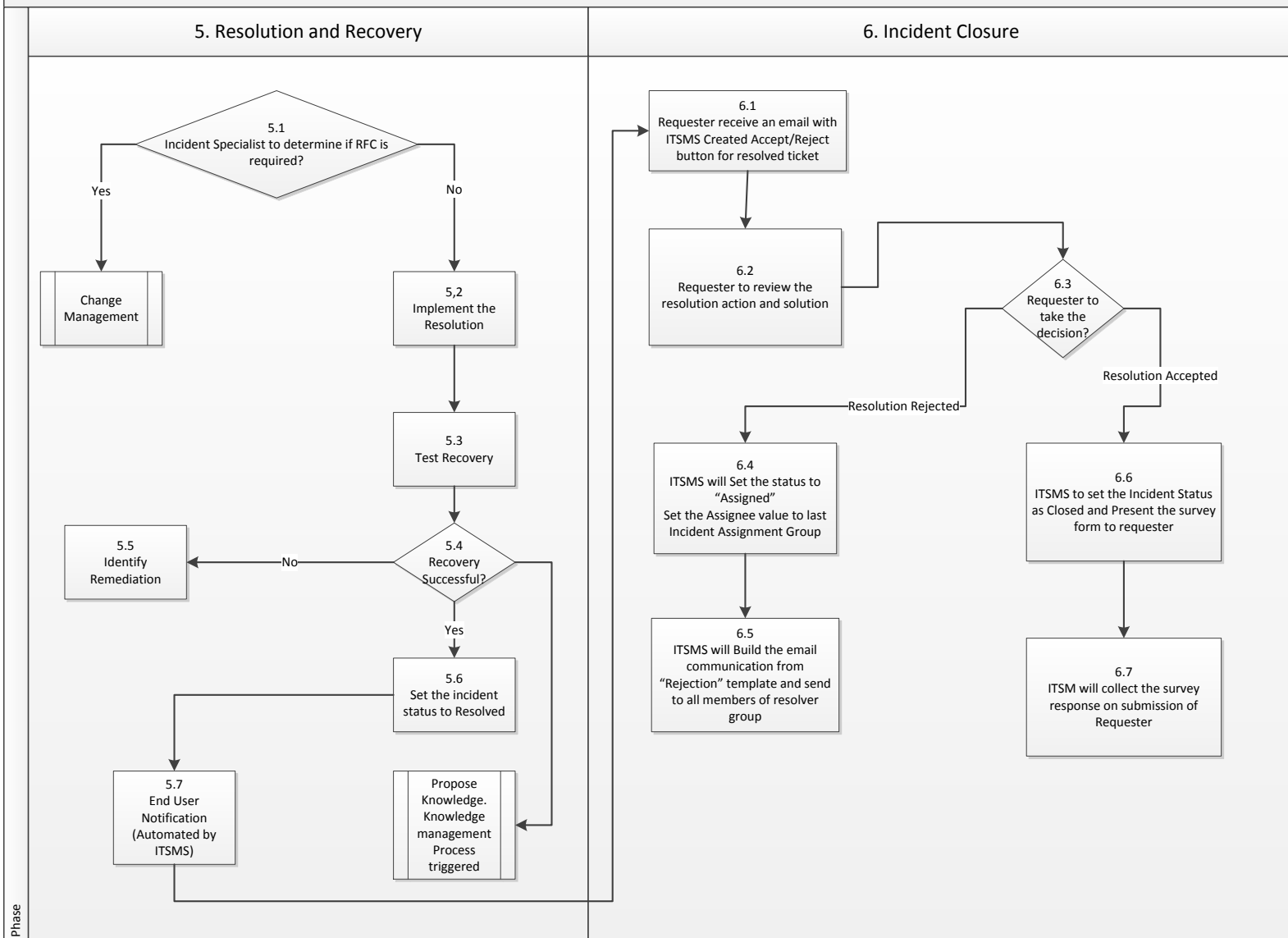
Příloha 4: SQL skript pro vytvoření dimenzionálního modelu

Příloha 5: Souhrnný SQL příkaz pro tabulku faktů

## Příloha 1: Incident Management Process

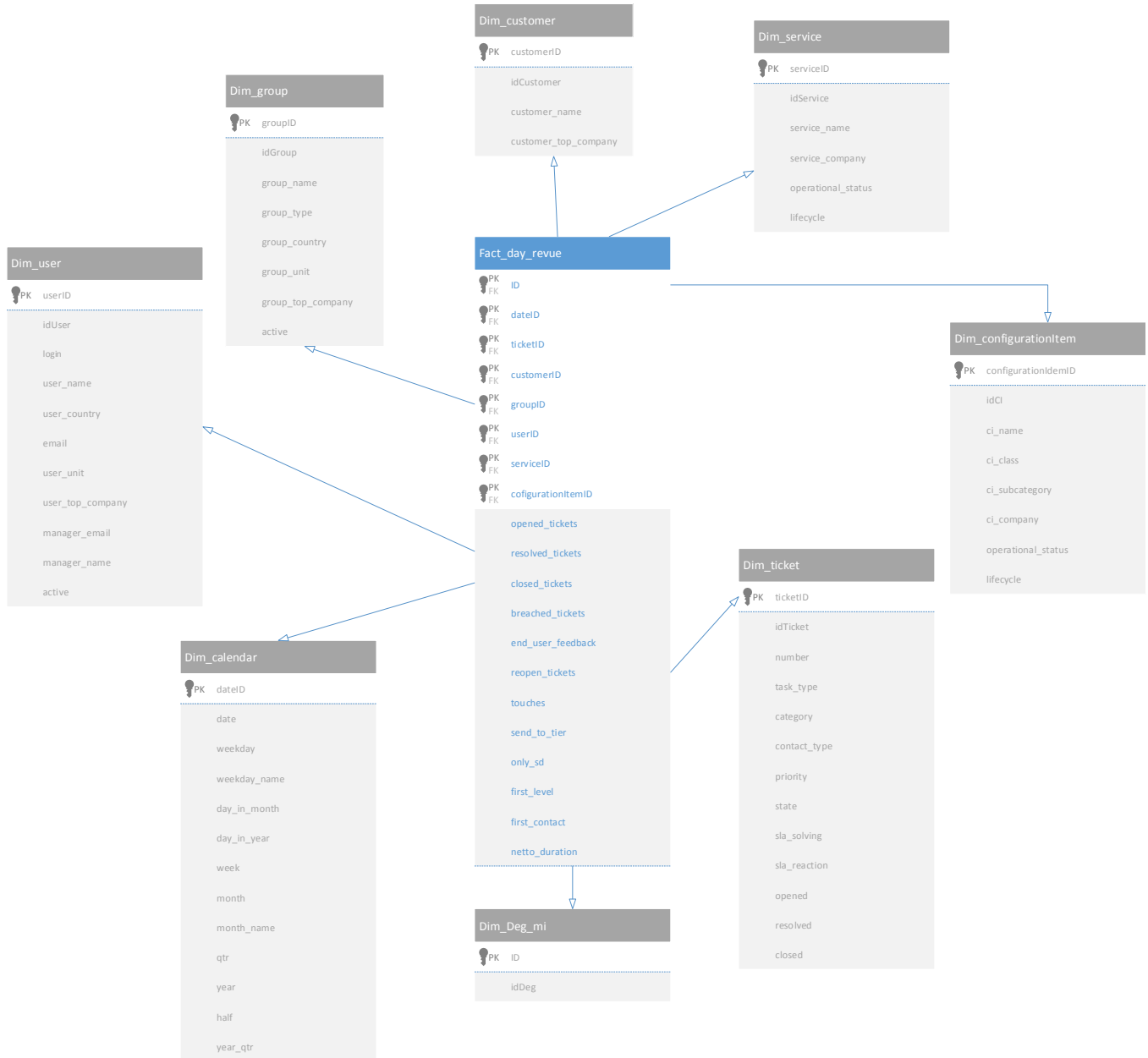


## Incident Management – Resolution and Recovery & Incident Closure

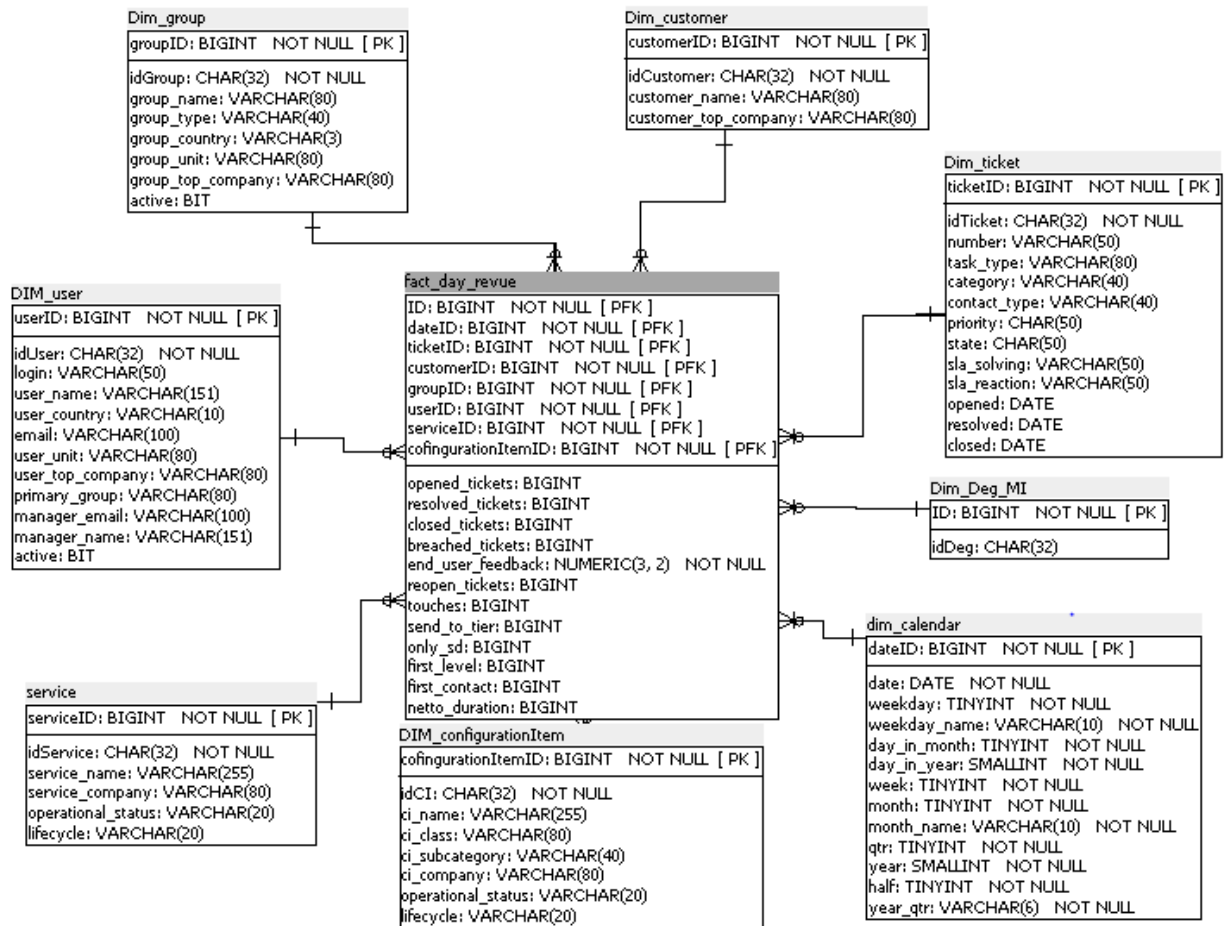




## Příloha 2: Logický model



### Příloha 3: Fyzický model



## Příloha 4: SQL skript pro vytvoření dimenzionálního modelu

```
CREATE TABLE Dim_Deg_MI (
    ID BIGINT IDENTITY NOT NULL,
    idDeg CHAR(32),
    CONSTRAINT DimDegMiID PRIMARY KEY (ID)
)
CREATE TABLE dim_calendar (
    dateID BIGINT IDENTITY NOT NULL,
    date DATETIME NOT NULL,
    weekday TINYINT NOT NULL,
    weekday_name VARCHAR(10) NOT NULL,
    day_in_month TINYINT NOT NULL,
    day_in_year SMALLINT NOT NULL,
    week TINYINT NOT NULL,
    month TINYINT NOT NULL,
    month_name VARCHAR(10) NOT NULL,
    qtr TINYINT NOT NULL,
    year SMALLINT NOT NULL,
    half TINYINT NOT NULL,
    year_qtr VARCHAR(6) NOT NULL,
    CONSTRAINT casID PRIMARY KEY (dateID)
)
CREATE NONCLUSTERED INDEX date
ON dim_calendar
( date )
CREATE NONCLUSTERED INDEX month
ON dim_calendar
( month )
CREATE NONCLUSTERED INDEX month_name
ON dim_calendar
( month_name )
CREATE NONCLUSTERED INDEX qtr
ON dim_calendar
( qtr )
CREATE NONCLUSTERED INDEX year
ON dim_calendar
( year )
CREATE TABLE Dim_ticket (
    ticketID BIGINT IDENTITY NOT NULL,
    idTicket CHAR(32) NOT NULL,
    number VARCHAR(50),
    task_type VARCHAR(80),
    category VARCHAR(40),
    contact_type VARCHAR(40),
    priority CHAR(50),
    state CHAR(50),
    sla_solving VARCHAR(50),
    sla_reaction VARCHAR(50),
    opened DATETIME,
    resolved DATETIME,
    closed DATETIME,
    CONSTRAINT ticketID PRIMARY KEY (ticketID)
)
CREATE NONCLUSTERED INDEX number
ON Dim_ticket
( number )
CREATE NONCLUSTERED INDEX opened
ON Dim_ticket
( opened )
CREATE NONCLUSTERED INDEX resolved
ON Dim_ticket
( resolved )
CREATE NONCLUSTERED INDEX closed
ON Dim_ticket
( closed )
CREATE TABLE DIM_configurationItem (
    cofingurationItemID BIGINT IDENTITY NOT NULL,
    idCI CHAR(32) NOT NULL,
    ci_name VARCHAR(255),
```

```

        ci_class VARCHAR(80),
        ci_subcategory VARCHAR(40),
        ci_company VARCHAR(80),
        operational_status VARCHAR(20),
        lifecycle VARCHAR(20),
        CONSTRAINT configurationItemID PRIMARY KEY (cofingurationItemID)
    )
CREATE TABLE service (
    serviceID BIGINT IDENTITY NOT NULL,
    idService CHAR(32) NOT NULL,
    service_name VARCHAR(255),
    service_company VARCHAR(80),
    operational_status VARCHAR(20),
    lifecycle VARCHAR(20),
    CONSTRAINT serviceID PRIMARY KEY (serviceID)
)
CREATE TABLE Dim_customer (
    customerID BIGINT IDENTITY NOT NULL,
    idCustomer CHAR(32) NOT NULL,
    customer_name VARCHAR(80),
    customer_top_company VARCHAR(80),
    CONSTRAINT customerID PRIMARY KEY (customerID)
)
CREATE TABLE Dim_group (
    groupID BIGINT IDENTITY NOT NULL,
    idGroup CHAR(32) NOT NULL,
    group_name VARCHAR(80),
    group_type VARCHAR(40),
    group_country VARCHAR(3),
    group_unit VARCHAR(80),
    group_top_company VARCHAR(80),
    active BIT,
    CONSTRAINT groupID PRIMARY KEY (groupID)
)
CREATE NONCLUSTERED INDEX group_name
    ON Dim_group
    ( group_name )
CREATE TABLE DIM_user (
    userID BIGINT IDENTITY NOT NULL,
    idUser CHAR(32) NOT NULL,
    login VARCHAR(50),
    user_name VARCHAR(151),
    user_country VARCHAR(10),
    email VARCHAR(100),
    user_unit VARCHAR(80),
    user_top_company VARCHAR(80),
    primary_group VARCHAR(80),
    manager_email VARCHAR(100),
    manager_name VARCHAR(151),
    active BIT,
    CONSTRAINT userID PRIMARY KEY (userID)
)
CREATE NONCLUSTERED INDEX User_idx
    ON DIM_user
    ( user_name )
CREATE NONCLUSTERED INDEX pmGroup
    ON DIM_user
    ( primary_group )
CREATE NONCLUSTERED INDEX email
    ON DIM_user
    ( email )
CREATE NONCLUSTERED INDEX mngEmail
    ON DIM_user
    ( manager_email )
CREATE NONCLUSTERED INDEX mngName
    ON DIM_user
    ( manager_name )
CREATE TABLE fact_day_revue (
    ID BIGINT NOT NULL,
    dateID BIGINT NOT NULL,
    ticketID BIGINT NOT NULL,

```

```

        customerID BIGINT NOT NULL,
        groupID BIGINT NOT NULL,
        userID BIGINT NOT NULL,
        serviceID BIGINT NOT NULL,
        cofingurationItemID BIGINT NOT NULL,
        opened_tickets BIGINT,
        resolved_tickets BIGINT,
        closed_tickets BIGINT,
        breached_tickets BIGINT,
        end_user_feedback NUMERIC(3,2) NOT NULL,
        reopen_tickets BIGINT,
        touches BIGINT,
        send_to_tier BIGINT,
        only_sd BIGINT,
        first_level BIGINT,
        first_contact BIGINT,
        netto_duration BIGINT,
        CONSTRAINT day_revue_ID PRIMARY KEY (ID, dateID, ticketID, customerID, groupID,
userID, serviceID, cofingurationItemID)
)
ALTER TABLE fact_day_revue ADD CONSTRAINT Dim_Deg_MI_fact_day_revue_fk
FOREIGN KEY (ID)
REFERENCES Dim_Deg_MI (ID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT dim_calendar_day_revue_fk
FOREIGN KEY (dateID)
REFERENCES dim_calendar (dateID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT Dim_tiket_day_revue_fk
FOREIGN KEY (ticketID)
REFERENCES Dim_ticket (ticketID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT DIM_configurationItem_day_revue_fk
FOREIGN KEY (cofingurationItemID)
REFERENCES DIM_configurationItem (cofingurationItemID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT service_day_revue_fk
FOREIGN KEY (serviceID)
REFERENCES service (serviceID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT Dim_customer_day_revue_fk
FOREIGN KEY (customerID)
REFERENCES Dim_customer (customerID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT Dim_group_day_revue_fk
FOREIGN KEY (groupID)
REFERENCES Dim_group (groupID)
ON DELETE NO ACTION
ON UPDATE NO ACTION
ALTER TABLE fact_day_revue ADD CONSTRAINT DIM_user_day_revue_fk
FOREIGN KEY (userID)
REFERENCES DIM_user (userID)
ON DELETE NO ACTIONON UPDATE NO ACTION

```

## Příloha 5: Souhrnný SQL příkaz pro tabulku faktů

```
SELECT mi3.sys_id9 as ID, cast(metric_instance.sys_updated_on as date) as
updateDate,isnull(task.sys_id2,'') as ticketID, isnull(core_company.sys_id4,0) as companyID,
isnull(sys_user_group.sys_id,0) as groupID, isnull(sys_user.sys_id1,0) as userID,
isnull(service.sys_id3,0) as serviceID,isnull(ci.sys_id3,0) as ciID,

(CASE WHEN ((metric_instance.value <> (Lead(metric_instance.value)
OVER (ORDER BY task.number, metric_instance.start)))
or task.sys_id2<>(Lead(task.sys_id2)
OVER (ORDER BY task.number, metric_instance.start)))
and task.sys_created_on=metric_instance.start
THEN count(metric_instance.value) END) as Opened,

(CASE WHEN ((metric_instance.value <> (Lead(metric_instance.value)
OVER (ORDER BY task.number, metric_instance.start)))
or task.sys_id2<>(Lead(task.sys_id2)
OVER (ORDER BY task.number, metric_instance.start)))
and ((mi3.value='Resolved - Awaiting Verification')
or (metric_instance.start=task.sys_created_on
and mi3.value='Closed'))
THEN count(metric_instance.value) END) as Resolved,

(CASE WHEN mi2.value='Resolved' and mi3.value='Closed'
and mi.value<>lead(mi.value)
OVER (ORDER BY task.number, metric_instance.start)
THEN count(metric_instance.value) END) as Closed,

(CASE WHEN metric_instance.start<=task.final_resolution_sla
and metric_instance.endd>final_resolution_sla
and (metric_instance.value <> (Lead(metric_instance.value)
OVER (ORDER BY task.number, metric_instance.start)))
or task.sys_id2<>(Lead(task.sys_id2)
OVER (ORDER BY task.number, metric_instance.start)))
THEN count(metric_instance.value) END) as Breached,

(CASE WHEN ((metric_instance.value <> (Lead(metric_instance.value)
OVER (ORDER BY task.number, metric_instance.start)))
or task.sys_id2<>(Lead(task.sys_id2)
OVER (ORDER BY task.number, metric_instance.start)))
and ((mi3.value='Resolved - Awaiting Verification')
or (metric_instance.start=task.sys_created_on
and mi3.value='Closed'))
THEN (SELECT avg(answer_integer)
FROM task_survey
left join survey_response on task_survey.instance=survey_response.instance
WHERE task.sys_id2=task_survey.task) END) as EUF,

(SELECT count(reopen_count)
where mi2.value='Resolved' and mi3.value='Closed'
and task.reopen_count>0) as ReopenCount,

(CASE WHEN mi.value <> (Lead(mi.value)
OVER (ORDER BY task.number, metric_instance.start))
or (mi.value = (Lead(mi.value)
OVER (ORDER BY task.number, metric_instance.start))
and mi.sys_id9 <> (Lead(mi.sys_id9)
OVER (ORDER BY task.number, metric_instance.start)))
THEN count(mi.value) END) TouchedPerson,

(CASE WHEN metric_instance.value <> (Lead(metric_instance.value)
OVER (ORDER BY task.number, metric_instance.start))
or (metric_instance.value = (Lead(metric_instance.value)
OVER (ORDER BY task.number, metric_instance.start))
and metric_instance.sys_id9 <> (Lead(metric_instance.sys_id9)
OVER (ORDER BY task.number, metric_instance.start)))
then count(metric_instance.value) END) as TouchedGroup,

(CASE WHEN task.sys_created_on=metric_instance.start
and left(metric_instance.value,3)='SD_'
and mi3.field='assignment_group' and left(mi3.value,3)<>'SD_'
```

```

        and ((mi3.value <> (Lead(mi3.value)
        OVER (ORDER BY task.number, metric_instance.start)))
        or mi3.sys_id9<>(Lead(mi3.sys_id9)
        OVER (ORDER BY task.number, metric_instance.start)))
THEN count(metric_instance.value) END) as Sendto2,

(CASE WHEN left(metric_instance.value,3)='SD_'
        and mi3.value='Closed'
        and task.sys_created_on=metric_instance.start
        and ((mi3.value <> (Lead(mi3.value)
        OVER (ORDER BY task.number, metric_instance.start)))
        or mi3.sys_id9<>(Lead(mi3.sys_id9)
        OVER (ORDER BY task.number, metric_instance.start)))
THEN count(metric_instance.value) END) as OnlySD,

(CASE WHEN task.sys_created_on=metric_instance.start
        and left(metric_instance.value,3)='SD_'
        and ((mi3.field='assignment_group' and left(mi3.value,3)='SD_')
        or mi3.value='Closed')
        and ((mi3.value <> (Lead(mi3.value)
        OVER (ORDER BY task.number, metric_instance.start)))
        or mi3.sys_id9<>(Lead(mi3.sys_id9)
        OVER (ORDER BY task.number, metric_instance.start)))
THEN count(metric_instance.value) END) as FirstLVL,

(CASE WHEN left(metric_instance.value,3)='SD_' and mi3.value='Closed'
        and task.sys_created_on=metric_instance.start
        and ((mi3.value <> (Lead(mi3.value)
        OVER (ORDER BY task.number, metric_instance.start)))
        or mi3.sys_id9<>(Lead(mi3.sys_id9)
        OVER (ORDER BY task.number, metric_instance.start)))
        and (task.contact_type in ('chat', 'phone')
        and task.u_only_handled_by_sd = 1
        or task.contact_type in ('email', 'integration')
        and task.duration < 3600)
THEN count(metric_instance.value) end) as FirstContact,

(CASE WHEN mi2.value in ('Resolved - Awaiting Verification','Assigned',
        'Work in Progress')
THEN sum(mi2.duration) end) as NettoDuration

FROM task
left join metric_instance on task.sys_id2=metric_instance.id
left join metric_instance as mi on metric_instance.id=mi.id
        and mi.start>=metric_instance.start
        and mi.endd<=metric_instance.endd
        and mi.field='assigned_to'
        and mi.value<>''
left join metric_instance as mi2 on metric_instance.id=mi2.id
        and mi2.start>=metric_instance.start
        and mi2.endd<=metric_instance.endd
        and mi2.field='incident_state'
        and mi2.value<>''
left join metric_instance as mi3 on metric_instance.id=mi3.id
        and metric_instance.endd=mi3.start
        and mi3.field in ('incident_state','assignment_group')
left join sys_user on mi.field_value=sys_user.sys_id1
left join sys_user_group on metric_instance.field_value=sys_user_group.sys_id
left join core_company on task.company=core_company.sys_id4
left join cmdb_ci as ci on task.cmdb_ci=ci.sys_id3
left join cmdb_ci as service on task.service_offering=service.sys_id3

WHERE metric_instance.field ='assignment_group'
        and metric_instance.calculation_complete=1

```

```
GROUP BY sys_user_group.sys_id, cast(metric_instance.sys_updated_on as date),
task.sys_id2,ci.sys_id3, service.sys_id3, core_company.sys_id4, metric_instance.value,
sys_user.sys_id1, mi.value, task.number, task.sys_created_on, task.contact_type,
task.final_resolution_sla, mi2.value, metric_instance.start, metric_instance.endd,
metric_instance.sys_id9,
task.reopen_count, task.u_only_handled_by_sd, task.duration, mi.sys_id9,
mi3.field, mi3.value, mi3.sys_id9

ORDER BY task.number, metric_instance.start
```